JOURNAL OF CONSULTING PSYCHOLOGY

VOL. 13, No. 4
AUGUST
1 9 4 9

EDITED BY
LAURANCE F. SHAFFER

ASSOCIATE EDITORS: EDGAR A. DOLL · WILLIAM A. HUNT · E. LOWELL KELLY · MORRIS KRUGMAN
BERTHA M. LUCKEY · FRED McKINNEY · CATHARINE C. MILES · CARL R. ROGERS · R. N. SANFORD

CONTENTS

PSYCHOLOGICAL EXAMINATIONS OF EMINENT BIOLOGISTS — ANNE ROE — 225

A PROJECTIVE TEST OF MASCULINITY-FEMININITY — KATE FRANCK AND EPHRAIM ROSEN — 247

STATISTICAL TESTS OF RORSCHACH ASSUMPTIONS: ANALYSES OF DISCRETE RESPONSES — J. R. WITTENBORN — 257

VALIDITY OF THE RORSCHACH 8-9-10 PER CENT — B. R. SAPPENFIELD AND S. L. BUKER — 268

PROBLEMS OF COLLABORATION BETWEEN THE CLINIC CENTER AND THE PSYCHOLOGY DEPARTMENT — FLORENCE R. DAY — 272

WECHSLER-BELLEVUE PATTERNS IN SCHIZOPHRENIA — SOL L. GARFIELD — 279

RELIABILITY OF THE SHORTENED MMPI — J. D. HOLZBERG AND S. ALLESSI — 288

THE ROSENZWEIG P-F TEST AS A MEASURE OF FRUSTRATION RESPONSE — J. C. FRANKLIN AND JOSEF BROZEK — 293

THE GUILFORD-SHNEIDMAN-ZIMMERMAN INTEREST SURVEY — J. P. GUILFORD, E. S. SHNEIDMAN AND W. S. ZIMMERMAN — 302

NEW BOOKS AND TESTS — 307

AMERICAN PSYCHOLOGICAL ASSOCIATION INCORPORATED

JOURNAL OF CONSULTING PSYCHOLOGY

ESTABLISHED 1937

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

THE JOURNAL OF CONSULTING PSYCHOLOGY is devoted to the fields of clinical and consulting psychology, counseling and guidance. It publishes the results of research bearing on these areas of psychology; reports, discussions, and case studies in clinical psychology; theoretical contributions pertinent to psychological diagnosis and counseling; and articles on the professional practice of psychology for human welfare.

THE JOURNAL OF CONSULTING PSYCHOLOGY was founded in 1937 by the Association of Consulting Psychologists, and from 1938 to 1945 was the official organ of the American Association for Applied Psychology. With the amalgamation of the latter society with the American Psychological Association in 1945, the JOURNAL was continued as the official publication in its specialty.

EDITORIAL OFFICE

Manuscripts submitted for publication, books and tests for reviews, and all correspondence on editorial matters should be sent to:

DR. LAURANCE F. SHAFFER, *Editor*,
525 W. 120th St.,
New York 27, N. Y.

Contributions relevant to the field of specialization of the JOURNAL are invited. Manuscripts should be typed with wide margins, and double-spaced throughout, including titles, quotations, footnotes and references. The style of all articles, figures, tables, and references must follow the form given by John E. Anderson and Willard L. Valentine. The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345-376. Copy for figures and illustrations must be prepared for reproduction.

Proof will be submitted to the author shortly before publication, and an opportunity to order reprints will be given at that time. Changes made by the author on the proof are chargeable to him.

BUSINESS OFFICE

Communications concerning subscriptions, change of address, claims for nonreceipt of a number and other business matters should be sent to:

AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1515 Massachusetts Ave., N. W.,
Washington 5, D. C.

THE JOURNAL OF CONSULTING PSYCHOLOGY is published bi-monthly, in February, April, June, August, October, and December, the yearly volume comprising approximately 440 pages. The subscription per year is $5.00, foreign $5.50; single numbers $1.00.

Copyright 1949 by the American Psychological Association, Inc.

Printed at 23 West Colorado Ave., Colorado Springs, Colo., and entered as second-class matter at the Post Office at Colorado Springs, Colo. Acceptance for mailing at special rate of postage provided for in section 538, Act of February 25, 1925, authorized May 27, 1940.

Volume 13 AUGUST, 1949 No. 4

JOURNAL OF CONSULTING PSYCHOLOGY

PSYCHOLOGICAL EXAMINATIONS OF EMINENT BIOLOGISTS

BY ANNE ROE
NEW YORK, N. Y.

THIS paper gives the technical reports of Rorschach, TAT and a special verbal, spatial, mathematical test administered to 20 biologists. The men who served as subjects were chosen for their eminence in research. All of them have made significant contributions to understanding of basic, normal life processes. All were born and brought up in the United States and all are presently connected with universities or research institutions.

The tests were administered as part of a project to study the lives, personalities and work of these men in an attempt to isolate any common features in their personality structure which would set them off from other groups, and to investigate any relations that there might be between personality and vocational choice and performance. The study has been financed entirely by a grant from the United States Public Health Service (Research Grant MH-10), and will be continued during the coming year, with a similar study of physicists. Full details of selection of subjects, and results will be reported in a forthcoming monograph.

In addition to the clinical studies of the individual men, I have group Rorschachs of 196 members of 8 biological faculties in large universities and that of the National Museum which will also be reported elsewhere [4].

This group included 8 geneticists (4 each plant and animal) 4 botanists (plant physiology and taxonomy), 3 biochemists, 1 anatomist, 1 physiologist, 1 paleobotanist, 1 paleontologist, and 1 bacteriologist. That there are no zoologists other than those doing genetics primarily is a reflection of present trends in research. Selection of the men was by the following process. Members in biology of the National Academy of Sciences and of the American Philosophical Society were listed (omitting those over 60, foreign-born or now primarily in administration) together with a few other names of invited speakers at recent important symposia. These lists were submitted to a group of raters (Detlev Bronk, Hans Clarke, L. C. Dunn, W. J. Robbins, J. R. Schramm and G. G. Simpson) who rated them on a 3-point scale and added other names. The ratings were combined, and from these a list of 23 men was selected. Three of these declined to take part, and the others are reported here. Their average age is 51.2.

All of the men are married; only one has been divorced. Average age at marriage was 27.8, and they average 2.2 children each.

Most of them were able to go on to graduate work immediately following their B.A. work. Average age of receiving B.A. or B.S. was 21.8; Ph.D., Sc.D. or M.D. was 26.0. They hold an impressive number of honorary degrees and have received many prizes and

other forms of professional recognition. All but three are members of the National Academy of Sciences and/or the American Philosophical Society, our ranking scientific organizations.

It is of interest that of the 23 men asked to take part in the study, only three refused, although the time requested was a minimum of six hours. Actually, in most instances considerably more time was spent with each man, and many of them were gracious enough to invite me to their homes. Social contacts, particularly in the family setting, afford an invaluable check on interpretation.

The Rorschach was given and scored according to Klopfer's system, [1] and in addition Munroe's Inspection Technique [2] was utilized, this being, in my experience, the most useful means of dealing with these data for group research purposes. Responses during the test were taken down verbatim in shorthand; the inquiry was not always taken down verbatim.

At the beginning of the study, some variation in the TAT cards selected was made, but as the work progressed nine were settled upon and used for each man. These were, in the male series, 1, 2, 4 (introduced fairly late after late sexual development appeared to be common), 6, 7, 10, 13, 15, 11. For the most part these protocols were also recorded verbatim, although occasionally I fell behind a few who talk very rapidly and were difficult to slow down.

The Verbal, Spatial, Mathematical (VSM) test is one compiled for the purpose by Dr. William W. Turnbull of the College Entrance Examination Board. There was no test available which had sufficient ceiling for this group. In fact, it took two attempts to get a verbal section which would serve, although the spatial and mathematical sections are more than difficult enough. Whether they will prove so for the physicists next to be studied is another question. There are, of course, no norms for this test, but no standardized test has norms which would be applicable to this group.

A life history was obtained, and each subject's research discussed at some length before any test was administered. They were always given in the order Rorschach, TAT and VSM.

THE RORSCHACH DATA

Quantitative data[1] are presented in Tables I, II, III and IV, which will not be verbalized in any detail.

That the number of responses is a little low for this level of intelligence is in general, I think, a result of the fact that professionally these men have arrived, although not all of them feel completely secure. Although for the most part very cooperative, most of them did not feel constrained to exert themselves to any great extent. I met with very little skepticism, expressed or not, over the nature of the technique. (One of them neatly turned the tables on me by presenting me, at our second interview, with several nicely done inkblots and a demand that I tell him what they looked like.) Except where time did not permit, or some other difficulty arose, I was able to go over the test findings with each man. There was practically no dissension, and frequent confirmation, sometimes of a rather dramatic sort.

The most striking findings are the prevalence of shading and color shock and the restriction of M. In view of the

[1]Administration of the test to 20 was not very brilliant. It should have been repeated with a request for no anatomical responses. This was not done, because he was quite disturbed and annoyed that he could see nothing but pelves and I felt that the suggestion would have been rejected, and might have prejudiced my being able to administer the other two tests. I am probably excessively unwilling to press subjects under these circumstances. I am not on the ground long enough, (nor sufficiently experienced in therapy to feel secure about it) to cope with any severe anxiety. This subject, incidentally, had gone out of town (not on business) and forgotten his original appointment with me.

TABLE I

RORSCHACH INSPECTION-METHOD DATA FOR TWENTY BIOLOGISTS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Subject:		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
R		39	33	41	26	23	16	19	11	16	26	11	15	23	17	23	29	20	10	28	15
T/R			—	—					+		—								+		—
Refusal									√				√			√		√	√√√		√
Location	W							+				+(BV)		B					+	+	
	Dd															+	+	+			+
	S	+		+	+		+		+	+										+	
	Suc	11	1								1			1		1	11				
Content	P																				—
	O	+										B		+							B
	At	+	+								+			++		+					++
	Range			—	+									t			t				t
Form	F%			+			+														+
	F							V		(VE)	B										B
Shading	ShSh	√	√		√	√	√	√	√	√√		√		√√			√	√	√√		?
	FK,Fc																		--		
	c																				
	C'				+							+							+		
	K,k	++								+											
Movement	M		--	+			--	—(M)	—(M)	—	(M)	--	—(M)	--	—	(M)		r			--
	FM		++				—	++	+	—				+		—	+				—
	m		+		+																---
	SM						—														
Color	C Sh				√		√	√√	√		√√	√√	√√	√	√	√	√	√	√		
	FC		—						--			—			—				--		
	CF			—		—				—				—	—	—				—	
	C																				
	S C			—	+						+	—			--						
	C:M			---	+		++				+				--		—				---
ITS		8	10	10	8	2	10	9	10	8	9	11	5	13	8	8	8	5	13	3	19
T%		23	27	10	8	13	25	21	18	19	4	9	7	48	0	4	10	5	0	0	20
W%		20	21	20	57	26	47	79	45	56	39	73	52	43	41	22	24	50	70	60	56
F%		46	45	63	35	48	59	32	36	38	31	36	47	48	59	48	34	45	20	25	67
F+%		83	87	96	89	91	80	50	75	83	100	80	100	64	90	91	80	89	100	100	40

TABLE II

RORSCHACH DETERMINANTS USED BY TWENTY BIOLOGISTS

Subj.	R	M	FM	m	k	K	FK	F	Fc	c	C'	FC	CF	C	W	D	d	I	S	P	O	T
1	39	3	5	1	6			18	2			2	1		8	17	5	7	2	5	13	9
2	33	1	8	1				15	7			1			7	17	7	2		6	6	9
3	41	9	4					26				2			8	24	3	6		6	2	4
4	26	3	4	1				9	2			4	2		15	8		3		4	7	2
5	23	3	4					11	2			3			6	16		1		6	5	3
6	17	1						10	2			3	1		7	6	1	2		6	5	4
7	19	1	3		1			6	1		1	2	2		15	4				6	2	4
8	11	1	3					4	2	1			1		5	5		1		4	4	2
9	16	2	1		1	1		6	3			2			9	5		2		3	5	3
10	26	3	3	1		1		8	3	1		3	3		10	14	2			3	2	1
11	11		1	1			1	4	2			1	1		8	3				3	4	1
12	15	1	2		1			7	1			2	1		8	7				6	2	1
13	23	1	4	1				11	2			4			10	11		2		4	11	11
14	17	2	2					10	1		1	1			7	8	1	1		7	2	0
15	23	6	1					11	3			2			5	11	2	5		5	3	0
16	29	3	7	2				10	5			3			7	15	3	4		8	5	3
17	20	4	1				1	9	2			2	1		8	8		4		6	0	1
18	10	2	3		1			2		1		1			7	2		1		6	2	0
19	28	5	6	1	1			7	3		1	4			17	7		2	2	8	6	0
20	15							10	2			2	1		8	4		2		2	9	3

TABLE III

RORSCHACH DETERMINANTS (IN PER CENTS), TIME, AND TURNING OF CARDS BY TWENTY BIOLOGISTS

	R	W%	D%	Dr%	F%	F*+ %	F†+ %	A%	Last 3	Turning	T/ R	Ave. RT	RT Range	d%
1	39	20	44	18	46	83	85	36	36	no	—	7.9	5 - 14	13
2	33	21	51	6	45	87	85	45	39	no	9	7.4	5 - 17	21
3	41	20	58	15	63	96	98	49	36	no	25	9.6	3 - 25	7
4	26	57	31	12	35	89	81	29	42	on I	30	17.5	8 - 27	0
5	23	26	69	4	48	91	96	39	56	yes	—	19.1	4 - 64	0
6	17	47	35	12	59	80	88	47	41	yes	57	19.5	4 - 40	6
7	19	79	21	0	32	50	42	48	27	yes	45	11.1	3 - 29	0
8	11	45	45	9	36	75	73	55	36	VIII & IV	65	28.3	3 - 100	0
9	16	56	31	13	38	83	75	38	31	no	30	10.5	1 - 30	0
10	26	39	54	0	31	100	81	31	35	no	28	2.5	1 - 4	8
11	11	73	27	0	36	80	63	46	36	no	49	8.7	1 - 19	0
12	15	52	48	0	47	100	80	47	47	on IV	44	33.5	8 - 102	0
13	23	43	48	9	48	64	74	39	48	yes	31	11.9	1 - 45	0
14	17	41	47	6	59	90	94	59	35	IV, V	39	9.6	2 - 22	6
15	23	22	48	22	48	91	95	26	43	no	39	26.8	1 - 79	9
16	29	24	52	14	34	80	79	55	28	yes	31	13.9	4 - 45	10
17	20	40	40	20	45	89	90	50	40	yes	36	22.9	1 - 65	0
18	10	70	20	10	20	100	70	60	20	no	78	28.6‡	4 - 53	0
19	28	60	29	4	25	100	96	39	39	no	38	11.1	2 - 25	0
20	15	56	28	14	67	40	60	13	27	no	28	17.0	2 - 35	0

*F+% of form responses only
†F+% of all responses
‡On 8 cards only.

TABLE IV

CONTENT IN THE RORSCHACHS OF EMINENT BIOLOGISTS

Subj.	H	Hd	A	Ad	Aobj.	At.	Sex	Sel.	A. At	Obj.	Pl.	N	Geo.	Art Des.	Mask	Abs.	Snow Ice	Blood	Expl. Fire	Cl. Wind	Emb.	X-Ray	Other
1	4	2	8	6	1	2	1	6		1	2	5										1	Puddle
2	1	1	13	2	1	5	1	3	2	1			1			1							1
3	9	5	11	9	2			3		2													
4	3	1	5	3	1	1		3	1	1	3		2			1	1						
5	3	3	8	1	2	3					1	1			1								refuse
6	1		7	1	2	1			1					3									1
7	2		7	1	1			1	1	2	2			2									
8	1		6		1	1		1						1									
9	2		5	1	1	1	1	1			2		1							1			
10	3		6	2	4	1	3		1						1	1		2		1	1		
11			5		1	2		1				1			1								
12	3		7		2			2			1											fountain	
13	1		9			7		4	1													1	
14	1	1	6	4	3				1		1												
15	4	3	4	2	2	4	1	2		1	1												
16	2	1	10	6	1	2		3	1						1				2				
17	3		10		2			1		1		1	1										
18	1		6		1			2								1				1			
19	3	1	10		2	1	1			3	5	1									1		
20			2		1	10		1									1						

importance of the former, I have listed below the responses of all of the men to Card VI on which shading shock is usually clearest, when it occurs. I am inclined to think that eventually a distinction will be made between shock determined by grayness or blackness and shock determined by tone variations, or texture. The latter is the commoner form in this group.[2] Material from the inquiry is indented.

VI. 2″ 1. Another peculiar kind of animal. I suppose it's the bilateral symmetry that makes me think of animals. Plants don't have it, nor do mountains. An impossible kind of monster, head like a caterpillar.

2. The rest is sort of amorphous, like a rug.

I suppose animal if torn up.

3. Lower down if you will pardon my crudeness, testes.

VI. It seems to me almost all are animal rather than plant life. I don't know what that looks like.

64″ 1. Looks like an insect emerging from a skin on which it has perhaps been feeding. Maybe it hatched there and grew to maturity and is coming out. There isn't any insect of that kind, but I —

Belonging to the cat family (the skin)

VI. 14″ 1. pretty . . . . the front part looks like a very peculiar flying animal of some sort.

(Front?) Going in that direction, the wind alignment I guess.

2. The main part is like a rug, fur, animal skin rug.

Shape rather than anything, the legs and — sprawled out.

3. Rather good face on each side half way down with a projecting set of chin whiskers.

4. Top of middle part, two little embryos, heads and trunks looking at each other.

VI. 10″ 1. Well that part here looks to me like the larvae that's been dissected, various parts of it but the rest of it doesn't come in at all, this part down here, just a lot of ink.

Such as a Drosophila larva. As marked. Of course the shading there gives you the

[2] C′ was never used as a determinant and FC′ very rarely as a main determinant. In these men FC′ seems to be associated with a certain amount of passive stubbornness and opposition which does not get expressed in S responses.

. . . . it suggests that different organs are present in the larva.

VI. 8″ 1. Now this, the front part reminded me > < (slowly) ∧ somewhat of a flying fish, the shape, 2 of them opposed, whereas this I haven't diagnosed, it makes it difficult. . . .

It even has eyes. as fluttering, they don't normally hold their fins out otherwise. It's not a good image, especially the wings, they aren't tattered like that.

2. Because it looks to me so like a sacrum or that part of the bony structure. The wings are too broad. I think I'll let it go at that. (Starts to put card down but pulls it back.)

3. There are some flat worms that look something like that; some quite broad flat worms that are blotchy but not with a head like that or such pronounced appendages on the side.

VI. 12″ 1. A sea robin escaping from a patch of sea weed.

A sea robin is a fish, rather good of it, it has wings and legs.

2. Also looks like a skin, very much, again upside down, the tail up.

It looks soft.

I don't seem to have so many ideas.

VI. 5″ 1. A hide pinned out. A wolf pelt.

Wolf partly because we had been talking of them. The shading, too, and the spread legs.

VI. 24″ 1. This part is certainly the skin of some kind of an animal, sheep skin for want of something better and . . . .

Sheep? Association with one I have on the blanket roll which is more or less that shape but is still not so dark as that.

2. This is an insect which I can't name. The larval form of a mosquito or something, an aquatic type. (Wings?) I don't think they have wings while in the water. Took the whole thing together with a large grain of salt.

VI. 17″ 1. Again like a stretched skin except for that uppermost part, or a rug lying flat, from an animal skin.

Clear. It felt like the fur side.

2. Upper part like the head of a tape worm a little bit.

Nothing in the rest.

a. There could be two heads facing each other, might be worm or caterpillar or something of that sort. The whiskers would be more in place then.

VI. 7″ 1. That's the skin of an animal. Whiskers. Has 4 legs. I don't know what kind, I never saw that kind, but some kind spread out. I suppose it could be used as a rug but would make rather a queer one but spread out in a rug fashion.

I don't think I would have thought of it as furry if I hadn't been struck by the shape in the first place. Wing effect is some kind of fur coming out of head.

VI. 20″ ∂ ∧ 1. A coon skin stretched out on a barn door drying. He's got a funny head but . . . .

It's tacked out or it's been tacked out and been dried in position, nailed at the stretched out points. It's the shading a little bit, this pattern in here isn't quite right for a coon skin but a definite pattern that suggests a hide.

VI. 5″ 1. A stuffed skin of leopard, or . . . . calls back African Hall at Museum.

I mean flat, not stuffed. It's stretched out, pulled. I don't know why I said leopard, some exotic animal.

2. Head of some sort of very indeterminate creature.

3. Contour map with a valley running down it.

4. A cleared specimen in oil of wintergreen to show the vertebral column. What annoys me is all this stereotyping. Wintergreen because that's used to make the specimen transparent.

VI. 40″ 1. Could be the beginning of a medical drawing but it isn't very meaningful.

Just a vague form. You have differentiation up in here, there is something different both in shape and quality, and a little bit of a thing here in the center (Couldn't be more specific about what it seemed to him to be)

VI. 7″ 1. This has much more of a face suggestion, a cat with whiskers, but also has insect relationships.

This diaphonous material is part of what suggested an insect, like a butterfly wing.

2. Primary node and streak of an embryo. Somewhat inaccurate.

3. Spread of a puddle; flow of something

spilled as though it was ink.

(Not including upper D)

VI. 29″ 1. Well, the principal part suggests somehow a suture that might have been in the body of some animal that was laid open but that isn't a very strong suggestion . . . . (pause of 68″). That's the only thing that occurs to me as though it might be a suture and some kind of vein or gut or something of that sort exposed.

(Cutoff W). It gives a sense of depth as though there was a trough there, there is some perspective.

V a. This suggests possibly some of those sticks that the African medicine men carry in their hands.

VI. 5″ A 1. Having spoken of a chick embryo, you get them everywhere with the neural plate at the end, in this case sprouting wings.

It's the axis and the differentiated region laterally.

2. Oh, I see, that looks like some sort of bear skin leaving out that front part here, spread out on the floor.

(Fc clear in inquiry)

3. Some sort of a jet-propelled missile going out here shooting out sparks laterally, suggested of something shooting out here from some sort of a take-off, you have the impression of a trough, but I don't know just what.

Something going out, not quite an animal, running along a trough spurting flame laterally. It suggests flame spurting out of something.

VI. Oh, heavens.

20″ 1. A symbol that might be used by a Navajo to represent in a schematic way a growing corn plant or perhaps a highly schematic animal in a Navajo sand painting.

W. The symmetry of the design reminds me of the sand painting. Would have to be very schematic. It's the color and shading.

2. The top of it looks like an insect head with bristles, the rest of it doesn't remind me of anything.

Upper D. Side pieces could be wings or possibly ?

VI. 7″ It doesn't look like anything . . . .

30″ 1. Something with whiskers on and

Upper D. Can't specify animal, just something but sort of head here . . . .

2. This looks like a vagina but it doesn't fit with that.

3. This is like a piece of the United States flopped over or something but it very strongly suggests a pattern.

Flopped because it didn't fit any known coast line.

VI. 25″ Well that doesn't look like much of anything I ever saw before . . . . 55″ that has me stumped, I'm just trying to think what it could look like. Trying to make something out of the . . . . something with whiskers on.

100″ I think that's a new species of (?) of illustration. I just see it as a *nice* symmetrical figure with out very much meaning. You can see a resemblance between certain parts and certain things but take the thing as a whole it doesn't look like anything I'm familiar with.

On inquiry said readily he had seen several things in it, and gave them as
a. head of tarsius, and
> b. Animal. A rigid pose, could be a lamb except for the face. See chart for both.

VI. 15″ Well, now this doesn't remind me of anything. Are you interested in knowing how it makes me feel? 40″ I never saw anything that looked like this, it has the old steer hide look, looking at this much but all the rest of this it doesn't look like anything at all; it doesn't remind me of anything.

Fur because of color, dark along the back, the color tone. Fc.

2. Except the designs you see on wall paper or linen, I never saw one like this of course.

V 3. Actually these look like the ovaries of a bird that makes it sort of a bird, beak, tails and feathers, but that's reaching for it, the answer is it doesn't look like anything.

If these were colored I might feel more, or have more answers. (This in response to inquiry into how it made him feel. Could not elicit any statement of feeling tone in reaction to the card.)

Mild degrees of color shock are not uncommon in normal groups. It is somewhat less prevalent than the shading shock and I think of considera-

TABLE V
ANALYSIS OF WHOLE RESPONSES IN THE RORSCHACHS OF BIOLOGISTS

Subj.	W	W%	Cards	Level			
				Combinatorial	Better than Popular	Popular or equivalent	Poor or vague forms
1	7	20	I,II,V,VI,VII	1	2	3	1
2	7	21	I-VI incl.	1	1	4	1
3	8	20	I-VI incl.	2	2	4	
4	15	57	all but X	2	6	3	4
5	6	26	I-VI incl.	2	2	2	
6	8	44	all but II & X		2	4	2
7	15	79	all		3	7	5
8	5	45	I,II,V,VI,VII		1	2	2
9	9	56	all but VI, VIII	3	3	2	1
10	10	39	all but IX	3	2	5	
11	8	73	all but III,VI		4	1	4
12	8	52	I-VII incl.		1	5	2
13	10	43	all but X		1	4	5
14	7	41	I-VI		2	4	7
15	5	22	I,II,III,VI,VII		3	2	
16	7	24	I,II,IV,V,VI,IX		2	4	1
17	8	40	I-VII	1	2	4	1
18	7	70	all but IV,IX,X	1	1	2	3
19	17	63	all but X	6	2	8	1
20	7	56	all but I,VII,X			2	5

bly less importance unless strong.

In their research it is difficult to argue that these men are not creative. Their major contributions have been, for the most part, too significant to be judged otherwise. Nevertheless, as was the case with the artists I studied, *M* responses may be very few, and when adequate in number are often remote in content or restricted in activity. This is almost as often associated with a corresponding lack of *FM* as it is with an excess. The *FM* responses also tend to be very mild.

Some other points are of interest. *W%* is not particularly high, although above average; excess of unusual details (almost always *dr*) is much commoner, although there are four protocols with none of these. The nature of *W* in this group is particularly interesting, and an analysis is given in Table V.

One subject gave many more combinatorial *W*'s than did any of the other men. His situation is a rather curious one. It is difficult for me to assess his work in this respect, but he does not appear to have very broad interests. On the other hand (and he is the only one in the group who makes a consistent practice of this) he invariably starts a paper with a statement of the general background and a very lucid exposition of how this particular piece of research fits into it which enormously simplifies the reader's task. He stands apart from the group in a number of respects. For one thing he seems to have become a scientist because he couldn't help himself, having the sort of mind he has, with really extraordinary powers of organization, rather than because science met an inner need, as seems to be the case with most of the others.

It is rather presumptuous of me to rate these men in terms of the breadth of their abstractions and theories, but I have made an attempt in all humility in order to try to check this against some Rorschach data. If the group is

divided into two parts on the basis of a rough Rorschach measure, constructed of the number of *W* better than *P* but not elaborate combinatorial, plus twice the number of the latter, plus one point for *M* over 3, and compared with a rough division on the basis of more or less breadth of abstraction, only two men are differently placed on the two bases. One is a man whose emotional difficulties are likely adversely to affect his judgment and the other is one who does much less generalizing than he might (perhaps because he is so very well balanced and controlled a man). This is very crude, and there are so many sources of error that it is included with misgivings, but only to suggest an approach to analysis of these relationships. What is perhaps more significant is that I was able for three of the last men I saw, after reading their work and before seeing them, to make not too bad predictions of the percent and quality of the *W* responses which would be given in the Rorschach. Again, this is only suggestive. One should probably also consider combinatorial *Dr*. These are not common, but probably have a similar meaning, perhaps with a special significance for limitation.

Technical responses (*T*) are rather few in this group, and in general more are given by anatomists than the others as might have been expected. Some of them, e.g. pollinium of a milkweed to D 12 of Card X were given by several subjects. Responses scored *sci* are not always *T*, the latter designation being given only when technical knowledge is clear.

F% is not generally excessive, but it should also be noted that responses in which *F* is not the dominant determinant are very rare. All of the *C'*, e.g. are *FC'*. The *m* and the *k* are more often *Fm* and *Fk* than *mF* and *kF*. For the most part *F+%* is high. Qualitatively, also, there are indications of greater than usual care with form in comments on inadequacies or inaccuracies, even though they may have been accepted. Also it was not uncommon during the inquiry to have a response recalled that had been rejected because it didn't quite come up to acceptable accuracy of form, e.g. two of these men considered but decided against frogs in the lower detail of VIII on this ground.

These men have much the same unaggressive, rather restricted sexual development and limitations of its expression that appeared in the study of artists, and that seems to be rather typical of professional groups in our culture. There is one outstanding exception, who can quite accurately be termed a "wolf" and I should think a very successful one, but this is probably the obverse of the same basic situation. One of the others insisted he had never been shy with girls but I am not too sure of this. The rest all gave, or admitted to, if asked, a history of social awkwardness, shyness with girls, and very frequently no dating to speak of until late college or graduate school days. An interesting additional point is the very common failure to give any response for the lower *D* in cards VII and VIII.

There is a considerable element of egocentricity in many of them which is determined more by disinterest in other persons, social uncertainties and lack of warmth, which are characteristic of many of them, than by any aggressive introversiveness, although it may sometimes have the appearance of ruthlessness.

THE TAT RESULTS

After experimenting with several "systems", I finally scored the TAT protocols following Wyatt [5], with only a few changes or additions. The basic data are given in Tables VI to X, which,

TABLE VI

SUMMARY OF ANALYSIS OF THEMATIC APPERCEPTION TEST PROTOCOLS OF BIOLOGISTS

Subj.	Story				Perception				Time Trend			
	S	SD	DS	D	Omit	Distort	Descrip. Detail	Deviation	Past, Present Future	Past and Present	Present and Future	Present only
1	6	7,13	2,10	15,11		2 church	(1) accident	13 (11)				2,6,7,10,13,15
2	6	1,2,4,10, 15,11	7			10 sexes	(1) (2)	1,10,11			1,2,6,7,10,15	4,11
3	2,6,7,10, 13							(7) (10) 13			6,7,10,13	2
4	4,7,10,13, 15,11	1,2,6					2,4	(2) (4) (6),7,(13)		7	1,4,6,10,13	2,11,15
5	1,4,6,7,15	2,10,13,11						(1) 4,10	13	4b,6,15	1,2,10,11	7
6	4,6,7	2	13	1,10,15,11			(1) 7,13	4 (7)				All
7	1,2,6,7,10	4,13,11		15	2		(1) (7)	(1) (7)	1,6		2,4,11	7,10,13
8	2,6,7,10 13,15,11	1			2	10, sex		(2) 10	2,10	13	1,7,15,11	6
9	10,13,15	6,7,11	1,2,4		1		1,2,4	(1) (4) (10) (13)	2,13	1,15	4,6,7,10,11	
10	1,4,6,7,10, 13,15,11		2					(15) (11) 6 (6) (13) (15) 11	6,7,10	15	1,2,4	13,11
11	1,6,10	2,4,7,13		15,11	1			(2) 6a (10) (15)		6a	1,2,6b,7	4,10,13,15
12	1,2,4,6,7, 10,13,15		11		1,4		11	(13)	4	2,13	1,6,15,10	7,11
13	1,2,6,7,10, 13,15	11						(1) (2) (6) (7) (10) 13,15,11	1,6,7,10,13,15,11		2	
14	4,6,7,10, 13,15	1,2	11		1,2,4			(6a) (13)			1,2,4,6,7,10,15 11	13
15	2,7,10,13, 15,11	1,6			1,13		6,10,13,15	(2) (6a, b) (7) (10) 15	1,10,15		2,7	6,13,11
16	1,2,4,6,7 10,13,15		11		1			(4) (7) (15)			4,6,15	1,2,7,10,13,11
17	1,4,6,7,15	13	2,10,11		1	10 sex	13	2 (4) 7, 10 (13)	1		4,6,7,11	2,10,13,15
18	1,2,6,13, 15,11,10	7			1			(7) (15) (11)			1,6,10,13,15,11	2,7
19	6,7,13,11		1,2,10	15				2 (6) 10			2,6,7,11	1,10,13,15
20	4,6		2,7,10,13	1,15,11					2		4	6,7,10,13

again, will not be extensively verbalized. Entries in the table are the numbers of the cards.

Description was resorted to chiefly for Card 11. The vagueness of the card was unpleasant and rather disturbing for many of these men who have a decided preference for the nonamorphous and factual. It has been noted in connection with subject 20 that he showed considerable disturbance. His TAT is largely descriptive, but nevertheless indicated that the depression he was suffering was not psychotic. Subject 6 thought it was all the "bunk"; although he went through the motions, he was basically noncompliant.

The omitted detail, listed under Perception, was most often the sheet of music in card 1. Distortions were: Subject 1, interpretation of barn in card 2 as church; Subject 2, reversal of usual specification of sex in card 10, and a woman leading a boy into evil; Subject 8, card 10 taken as mother and daughter, whereas Subject 17 took it as a monk and a boy.

Entries in parentheses in the column head *Deviation* refer to stories which are not uncommon themselves but have an unusual twist. Other entries are for unusual stories. It should be stated that the reference point is my own experience, which is pretty well limited to superior normal adults.

The time trend is extremely curtailed, more so for the past than for the future.

Level is dominantly concrete-factual;

TABLE VII A

SUMMARY OF ANALYSIS OF THEMATIC APPERCEPTION TEST PROTOCOLS OF BIOLOGISTS

(Continued)

Subj.	Level					
	Concrete factual	Endopsychic	Symbolic	Mythical	Make-believe	Conditional
1	6	2,7,10,13	15		11	
2	1	2,6,7,10,15 (1)	11			4 (1)
3	2,6,7,10,13	(10)				
4	1,2,4,7,10, 13,11 (6)	6 (1) (15)	(7), (11)	15		
5	2,4,6,7,10,13	1,15 (13)			11	(1),(4),(7)
6	1,2,4,6		15			7,10,13,11
7	6,7 (13)	1,4,10,13	2,15?	11		(1)
8	7,13	1,2,6,10,15 (13)	(15)	11		(1)
9	1,2,6,7,10, 13,11 (4)	4,15				6,13
10	1,2,4,6,7,10, 13	15 (1),(4),(10)	11			(1)
11	1,2,4,6,7,10, 13	(1) (10)	15			(6), (7)
12	2 (1), (4)	1,4,6,7,10,13, 15		11		
13	1,2,6,7,10, 13,15	(1),(2),(3),(10), (13)			11	(2),(6)
14	1,2,4,6,7,10, 13,15	(1)			11	(6)
15	1,2,6,7,10,13, 15	(1), (10)			D11	(6)
16	2,4,6,7,10,13, 15,11	1 (2),(6),(7), (15)				
17	1,2,4,6,7,10	13,15 (2),(7),(10)				11
18	6,7,10,15a (1)	1,15b	2 11?		11?	(13), (15)
19	2,10,13,11	1,6,7,15				(1), (10),(13)
20	2	1,4,6,7,10,13 (2)				(7)

TABLE VII B

SUMMARY OF ANALYSIS OF THEMATIC APPERCEPTION TEST PROTOCOLS OF BIOLOGISTS (Continued)

Subj.	Tone											
	Indifferent	Detached	Contemplative	Cheerful	Serene	Melodramatic	Unhappy	Tense	Anxious	Morbid	Aggressive	Sardonic
1		6	2,15?			13,11		7,10				(7) (15)
2				2			10,15	4,7,11	1,6			
3						7,13	6,10		2?			(13)
4					2	15	6,10,13	1	4,7		11	
5	1a					10,11	6,13,15	2,4a,7	1b		4b	
6				11		2,4,6,7 10,13,15		1				(15)
7			2		1	11	6,13	10,7	4			
8					7,10	11	13,15	1,2	6			
9							10,15	2,6,7	4,11	1? 13?		
10					2	11	4,10	6,7,13, 15	1			
11			2?, 15?			4	1,6,7 10,13					
12					(10)	15,11	4,6,7 10,13	1,2				(15)
13				1,2, 13,15		11		7,10	6			(15)
14			1			11	3,10,13, 15	2	6		7?	
15							1,2,6b, 10,13,15,11		6a		7?	
16							1,10,13	2,15,11		15?		
17	11					4,6	4,6,7 7,10,13	1,2				15 (4)
18					2,7,10		13,15	1,6 (7)	11	13a		
19	1					11		2,7,10, 15	6,13			
20	1,11			2			4,10,13	6,15	7			

TABLE VIII

SUMMARY OF ANALYSIS OF THEMATIC APPERCEPTION TEST PROTOCOLS OF BIOLOGISTS (Continued)

Subj.	Quality							Personal Relations				Presses				
	Simple→ Literate	Simple	Literate	Cliché	Col	Cr	Ev.	Formal	Emotional	Not Stated	None	Friendly	Unfriendly	Impersonal	Internal	?
1			2,6,7,10, 13,15		(7)	(6) (7)		2,6,7	13	10	15	7		10,13	2,15	6
2	1	6	2,4,7,10, 15,11					1,2,6	7,10,11	4	15	2,6,7	10,11		15	1,4
3	2	10 (7)			6,7,13			6,10	2,7,13				7,13	2,6,10		
4	All	2						1,2,6,10, 15	7,13,11			1	11	6,10,13,15	4	2,7
5			All			(10)		6,13,15	2,4,7,10, 11		1	4a, 7b	2,3b,7a	13,15,11	1	6,10
6	11				1,2,4,6, 7,10,13,15	(1),(7), (10),(13)		2,6,7	2,4,10,15	13	1,11	4	2	7,13	1,10,15	11
7	10	4,6,7	1,2,15,11			13(4)		6,7	4,10,13	2	1,15,11	4,6,7	11	13	1,10,15	2
8			All					1,2,6,7,10, 13		15	11	1,2,6,10	7	13,15,11		
9			All			(2)		2,10,15	2,4,6,7, 13		1,11	7	6	1,10,11	2,4,13,15	
10			All					1,2,4,6,7, 10	13,15		11	1,4,6,7	15	2,10	13	11
11	7,10,15	1,2,6,13			(6)	(2),(4), (6),(7)		1,2,6b,10, 13	4,6a,7		15,11	1,4,7		6a,10,13,15	2	11
12	1,2,4,13, 15,11	6,7,10						1,2,6,7	4,10,13,11		11	1,6,7	11 6,11		2,4,10, 13,15	
13			All					2,6,7,10, 11	1,13,15			1,2,7, 13,15		10		
14	1,2,15,11	4,6,7, 10,13				(7)		2,6b,7,10, 13,15	4,6a		1,11		6a,7	2,10,13,15,11	1,4	
15	1,2	7	6,10,13, 15,11	(7)		(7)		1,2,6a,10, 13,15	2,6b,7		11	1	7	2,6b,10,13,15		11
16	1,2,4,15	6,7,10, 13,11						6,7,13	2,4	10	1,15,11		2	6,7,10,13,11	1,15	
17	15	1,2,4,10 13,11		6,7 (4)				1,6	2,4,7,10, 15	13	11	1,10	2,4	6,7,13	15	11
18	All							2,6,7,10	13		1,15,11	6,7,10	13	15,11	1,2	
19	6	1,10,13, 11	2	(7)		(2)	(15)	6	2,7,10,13	1,15	11	1	2,7,10	6,11	13,15	
20	All							2,6,10,13	4	7	1,15,11	4 7,6,7		2,10,13		1,15,11

TABLE IX

SUMMARY OF ANALYSIS OF THEMATIC APPERCEPTION TEST PROTOCOLS OF BIOLOGISTS (Concluded)

Subj.	Outcome						Certainty of Outcome			Personal Reference	Refused	Personal Opinion
	None Possible	Success	Defeat	Unsolved Tension	By rule	None or ? given	Certain	Probable	Possible			
1	15,11					2,6,7,10,13					1 (11)	
2		1,2,6,7	10,15 ?			4,11	1,2,6,7, 10,15			1,2,6,7		
3		13	7,10			2,6	13	7,10		10 ?		6,7
4		4,10,11	1	6	13	2,7,15	10,11	1,4,6,13		1		4,7
5		10,13 ?,4b	11	2		1,4a,6,7,15		4b,10,13, 11	2			(4)7
6	1,6,7,10, 13,15,11					2				All but 15		
7	7	1,2,4 man 11	4 girl 6 11 others			10,13,15	2,6	4,11	1	7		4,13
8		2,7,10 for 1	11	15		6	10	2,7,11, 15				
9		2,11	13	6,7	10	15,1		2,4,6,7	1,10,11,13			2,6,7
10		1,2,10,11	4,7,13	6		13,11,15	6	1,4,7,10	2			
11	6a,13,15	7,2	1,6b			10,11,4	1,2	7,6b		1,7,10,15	(11)	2,4,6
12		6,15		1,4	10	2,7,13,15,11	1,6,15	4,10				
13		1,2,7,10,13, 11		6 (7), (10)		15	1,6,10, 13,15,11		2	7		
14	13	1,6,7,10	4,11	2,15		(13)	1,2,6	4,5,11	7,10	1		7
15	6b,13	1		2,10,15	7	6a,11	15,2	10	1,7			
16	13	4	6	15		1,2,7,10	6	4	15			
17	13,15	4,6,7,11		1,10		2	6	4,7,11				
18	2,7	1,6,10,11	15	13	11		13,11	1,6,10	15			
19			7,11	2	6 (2)	1,10,15,13	6	2,7,13, 11				2,7,10,13 11
20	1,13,15	2,4				6,7,10,11	4	2				

the exceptions are men who are presently most concerned with specific emotional problems. Entries preceded by a plus have the meaning of additional scores on the Rorschach.

Tone is dominantly unhappy and tense. Melodramatic stories are usually limited to cards 11 and 15, and reflect their way of handling stimuli which seem to them rather unrealistic. This group for the most part, draws a very sharp line between fact and fancy, and they take a rather caustic view of the latter.

Quality is chiefly simple to literate.

Personal relations are somewhat oftener formal than emotional; presses mostly impersonal. I have introduced the category "internal" here to take care of situations in which the protagonist is impelled by inner needs.

Outcomes that are given are more often successful than not, but for many in which no outcome is given a continuation of tension is implied. Degree of certainty is usually not great.

Cards invoking specific personal references or opinions of the characters or their behavior are also noted.

Themas for each card have not been listed. Instead it has seemed more useful to give all of the reponses to one card, and I have selected card 10 for this purpose. These are reproduced as they were taken down.

Qualitatively, outstanding aspects of the protocols are the generally nonaggressive attitudes, the mildness of heterosexual interests and feelings, and at the same time the importance of and security in marriage. Paternal attitudes are sympathetic. Some of the protocols have a strong feeling of determination and resulting success; but not all show this. There is, however, an almost total lack of any expectation of or reliance on luck, or fate, or any other nonpersonal intervention as an aid. The nonaggressive attitudes are firmly entrenched, but there is a hint of underlying resentment or rebelliousness, which, however, is generally attached to characters in early maturity, and which is usually not expressed in action and if it is, results unfortunately for the one expressing it. On the whole, there is a strong conventionality in most of the protocols, the only common break in this being in connection with religion, conventional forms of which most of them have discarded.

I should like to comment on the longest story reproduced below. Long, rather dramatic narratives were the rule in this protocol, which is completely different from all the others. Furthermore there were several stories which show this same repetition of the scene which is unique in my experience and of the significance of which I am not certain. It may be an indication of the repetition compulsion which is characteristic of many neurotics, although I do not have life history data of a sort to indicate that this is particularly strong with him. He has more, and more obvious difficulties than many of the men, a more irregular professional history, and certainly a much richer affective life.

Responses to Card 10.

10. I don't know why I think it's Russian, it's in the face of the woman and the eyes of the man, Central European. The age of the individual involved bespeaks an affection lasting late in life. That she isn't young is shown as skilfully by the way her hair is done as by her face. It is, however, a story of very firm irrefragable affection. Maybe he is a Russian general going off to war (?) I don't think the author need be concerned with that because he has no answer to that.

10. (Long pause.) My first reaction is a bit of resentment that the drawing is so bad I can't tell if this is a man or a woman. The first reaction is that this is ardently erotic, affectionate kissing, but it's more idealistic. I took at right a man, at left a woman, kissing brow rather than lips, a certain amount of childish-

ness in that R with rounded cheeks, the pose of the hand like a child's hand on the mother. But the left looks wicked. Something not right about this pose. This a question of evil is some way? This gives me much vaguer images than the other one. (?) This one on the left will take charge of that one and lead him into evil. Not a contemporary though the bad companion. Older woman taking charge of a childish boy and he's being misled and will fall for it because not strong enough to resist it.

10. 15". These are getting progressively no better . . . . Is this a graded series and they get worser? . . . . 1' 15" I don't get much of this one, well I'd say it was a husband and wife scene in a family with pretty complete confidence and considerable affection among the members although they have been married for a long time. The husband has been having physical difficulties or he has been ill or something and has just been told he has a disease there isn't much hope for and is breaking the news to her. She is broken up but bucking up and taking it in good style. She needs comforting from him. I haven't any future, I expect he is going to up and kick the bucket sometime fairly soon but this is a matter of months.

10. I think this is probably not a romantic scene because the man and woman involved are not very young any more. I think they are probably married and have been for some time and probably have a considerable family which she has to stay home and take care of still. The husband is departing on a long trip to foreign parts and possibly from the nature of his expression he might be an army officer and might be an explorer. In any case he's going away for a long trip and as he leaves his family he is embracing his wife and she is clinging to him and they aren't kissing each other with sensual enthusiasm, but they are just clinging to each other while they say good bye and I think the wife is feeling worse than the husband is. I think it's all going to come out all right. He'll come back home. I might say by way of explaining that last remark that I say that because he isn't young and whether an army officer or explorer he will not do anything very dangerous. If an army officer he won't get into anything dangerous because he is a big shot and if an explorer he has been around enough and he will come home perfectly whole.

10. (Long pause, turns the card around) That's a puzzler. Well we have here a lady who is respectable but not too determined and a gentleman who is old enough to know better. The lady has a comfortable nest egg laid away. The gentleman has learned about it and has reached the point of making a proposal to the lady that they marry and he is kissing her quite chastely on the forehead and she has laid her hand up on his chest, closed her eyes and I think is going to consent. The shadows are not quite right in this picture which makes it a little bit difficult. The shadows here. (Points out that the deep shadow on her face would indicate that the light is coming from a different direction than seems to be the case from the rest of the picture.)

10. That's quite a good drawing. The black background leaves an awful lot to the imagination the way the faces are cut off. Well, it could either be very chaste and respectable or heavy necking as far as I can see. A fond farewell. It need only involve the higher passions. I don't care much for a girl with such short fingers. If I were in such a clinch I'd rather have a different girl. But it doesn't matter much to them. (?) They don't seem particularly happy so that's why I think it wasn't the grand passion.

10. This gets worse and worse. These are only details. Well, these are two lovers that don't have much of an opportunity to see each other and now they have met and have sunk into oblivion. That's about all I can say for that. You could spin any number of yarns before and after but it isn't suggested by the drawing itself.

10. The lower figure is definitely a woman. The upper figure I think probably is, too. uh. . That's possibly, well, I'm not sure whether to call it a prodigal daughter, well, we'll call it that. Perhaps a daughter who has been away, has gotten into difficulty, has lived on the Bowery or that sort of environment and has finally come back home and mother has taken her to her arms and is forgiving her so that it is a reconciliation and forgiveness and all happy there ever after.

10. Um, oh, they are a married couple and something has happened, maybe her mother has died. I don't know what, anyway she has come to him for comfort. And I assume that time will lessen her distress. If you would like me to think up more lurid situations I can.

10. Hmm. Well, this is a husband and wife

who just heard of a very terrible tragedy that has happened to one of their children probably and they turn to each other. They probably have been married long enough and have gotten along and will share each other's grief. They probably will have other moments like this in their lives and will share it and keep getting closer to each other because of it.

10. Well, this I'd say was a goodbye, farewell scene. A man is leaving on a trip, maybe called into the service, mobilization, or something and that's his wife and she's just realized what a hell of a fine fellow he is after living with him all these years, so she's very heartbroken and of course he's very sad too about leaving his family. A very nice picture. Yes a touching scene. Reminds me of a picture which was in Life a while ago, a man coming home, it was the real thing. . . .(?) Pfft. A matter of probability, what are the odds of getting killed?

10. Oh, gosh, you are getting them more difficult, aren't you? Gee, you ought to give me some that aren't so sad, these all look sad to me. Well, this is a woman and her husband or equivalent, her man any way, and she is very sad about something and he's trying to comfort her. I don't know what she is sad about. It could be a million things and I don't know what is going to happen. I suppose she will get over it, people do. I can't say how they got into this situation or what the outcome is going to be.

10. It's a little hard to see just what's there, it's going to be a little difficult. This woman is such an indefinite age that I don't know whether it's meant to be his sweetheart, wife, mother . . . . or even sister. (His voice dropped on this.) You see he looks somewhat younger than she (is it she or her? I can't remember) and yet one can't see him very plainly either, so—(E: You can have it the way you want it.) Well all right, she's not quite the type I'd pick for a sweetheart. This looks like a much more commonplace story to me. One could imagine that . . . . that they are a married couple and that he has had some. . . .well at any rate it might be that he's got to go away somewhere and leave her and that they are pained at parting. His work probably is taking him away; we won't say it's war but any way some work that takes him far away and he might be for instance, be a . . . . it's hard to make it plausible I wanted to say that he had succeeded in getting a position on a ship that he long had been wanting to get and that they had been living together in town at some clerical work he didn't like and he had always longed to get away and get some work in which he had a real interest. He has perhaps been working in the office of a shipping company he hardly to me looks the type for an engineer. They look like rather commonplace people, that is commonplace occupations, and he has had a desire however to have a job on a ship as, let's say as purser, and has finally gotten this job but it's a ship that takes him to the orient and will take him away on long journeys. He's perhaps gone on short ones but it's painful parting because she has meanwhile become pregnant and he will be away for many months and the baby will be born in the meantime and she regards it as dangerous because there is danger of war breaking out and all ships will be in danger so they have mixed feelings. Both of them. He is glad to have the job but is sorry to be away when the baby is being born. She doesn't like that either and is worried about him being in danger. Then we pass on from the picture. He does go off from her on his ship. He writes to her but the letters take a long time. . and she to him. . and. . . . shortly before her baby is to be born war does break out and he is interned by the enemy let's say in Japan but she has no news of him and doesn't know what has happened to the ship. She has a terrible time in childbirth and doesn't know whether her child has a father or not, alive, but the son is born, she receives payment from the company and manages to get along and years pass by and she hasn't been able to get any news of him . . . . until one day before the war is over a wild looking man with a beard rings her door bell and she finally recognizes him. He managed to escape from internment and get off on a Chinese junk and through very exciting adventures around the world gets back, and then the question arises what he should do and she doesn't want him to go again but now he is resolved to go and after being a little while at home together he starts off again because the merchant marine is very important and there is another parting but this time he's determined to help . . . . to help his country is dominant in him and he, he has even a feeling of joy in going. Now we might perhaps leave at this point.

10. Trying to settle on the age, there, that's a main problem. Well I would guess that that is a, — that they aren't too young. It's presumably a husband and wife and there has

been some difficulty, shall we guess with the child? Possibly died or there has been some accident and they are comforting one another. As for the future I guess they will get over it.

10. (Long pause.) I would say that the events preceding the picture were these. This seems to me to be a young man here on the left and that he had been away from home for a long time either in some hazardous undertaking such as war which isn't indicated at all by the picture so far as I can see, or that he had not been heard from and it was feared that he was in bad company or getting into trouble one way or another. This picture shows him returning home with the great relief of his mother, I would say, greeting him while he still has a rather jolly, carefree expression as much as I can see of his face, so she's the one expressing emotion as far as I can see, he just shows a little awkwardness, so he's not so completely carefree. I would suspect if my guess about the earlier part is correct the outcome would not be so happy. He would be likely to get into the same situation again unless it was a case of war which is also likely again but not for the same reason.

10. I can't even be quite sure of the sex of the various people involved here. I suppose a man and woman here but not sure of that, first impression wasn't that. He's kissing her cheek. Well, I suppose you could say that they have heard bad news and he's trying to comfort her, something of that sort I'll let it go at that (firmly.)

10. Oh, that's tougher. (Long pause) Well I suppose they intentionally leave out enough so you can't be sure what it's intended to be. I'm going to guess that the left hand character is a monk or a priest and that he's trying to provide some comfort for this broken hearted lad on the right but that is a poor diagnosis, it doesn't fit much of anything. I don't know whether priests are in the habit of comforting young men who are in trouble or not, my experience being so limited.

10. Well there is a great deal of affection here. I presume this can be interpreted as a husband and wife, at least it would be a characteristic pose, the wife with her head on the shoulders of her husband. Whether happiness or grief they are certainly sharing the situation. (?) I should think favorable outcome, I should think those two would find strength in each other.

10. I don't think it's happened yet but I think it will before very long. I don't just see why they should have such a brilliant light. Might be a headlight of a car that arrived at an unfortunate time such things do happen. I don't think any one would stand there on purpose. Might even be one of these flashlight pictures that gentlemen with low minds take sometimes.

10. I would interpret that as grief stricken parents after the loss of a baby. They seem to be relatively young which would fit that idea. To me the attitude indicates they are trying to comfort each other.

THE VERBAL-SPATIAL-MATHEMATICAL TEST (VSM)

This test was compiled by the College Entrance Examination Board from several of the most difficult of their tests. The verbal section consists of two parts, comprising 79 items in all. The second part was added when it was apparent that the first did not have enough ceiling for this group. It had been hoped that this would be an adequate power test, but in some instances at least, it is primarily a speed test. The problem in each set is to find the antonyms; in the first part four words are given, from which the two antonyms must be selected, in the second part 6 words are given, and the task is to find which of the last five is opposite to the first. Time limit was 15 minutes.

The spatial section comprised 24 items, each item consisting of 4 views of solid figures. The task is to select the two views which are of the same figure. This particular test was suggested by the Board, because they believe it to be the purest space test they have. Time limit was 20 minutes.

For the mathematical section, 39 items from a 60-item test were selected so as to omit some of the items at easier levels of difficulty. The original test was given with a 60-minute time

TABLE X

NUMBER TRIED, RAW SCORE, AND SIGMA SCORE OF BIOLOGISTS ON A VERBAL SPATIAL MATHEMATICAL TEST

Subj.	Tried			Raw Score Right			Sigma Scores		
	V	S	M	V	S	M	V	S	M
1	71	24	22	50	6	11	—.54σ	—.77σ	—.95
2	79	22	18	73	10	12	+1.35	+.14	—.78
3	79	14	28	70	10	18	+1.10	+.14	+.24
4	79	19	25	68	15	22	+.94	+1.27	+.91
5	68	23	12	58	3	11	+.11	—1.45	—.95
6	60	24	22	51	7	8	—.46	—.54	—1.45
7	79	21	15	59	5	6	+.20	—1.00	—1.79
8	77	18	24	68	9	10	+.94	—.09	—1.11
9	79	11	22	61	9	21	+.36	—.09	+.74
10	79	24	27	71	20	18	+1.18	+2.41	+.24
11	—	—	—	—	—	—	—	—	—
12	70	24	25	50	15	22	—.54	+1.27	+.91
13	63	22	28	57	8	24	+.03	—.32	+1.25
14	79	22	25	65	18	19	+.69	+1.95	+.41
15	52	8	23	44	8	22	—1.03	—.32	+.91
16	77	10	30	69	7	27	+1.02	—.54	+1.76
17	55	10	20	43	8	14	—1.12	—.32	—.44
18	50	15	15	28	7	13	—2.34	—.54	—.61
19	69	12	21	52	9	19	—.38	—.09	+.41
20	60	24	21	38	4	18	—1.52	—1.22	+.24
Mean	69.8±2.3	18.2±1.3	22.3±1.1	56.6±2.8	9.4±1.0	16.6±1.4			
SD	9.89±1.6	5.63±0.9	4.63±0.8	12.15±1.9	4.42±0.7	5.94±1.0			

limit, but only 30 minutes was allowed for this administration and no subject attempted more than 30 items.

There are no norms for these tests, nor any comparative data. On the other hand, there are no standard tests which have enough ceiling for this group, nor any with norms which would be applicable to this group even if there were enough ceiling. The data were secured for information on individual differences within the group and for comparison with other groups of scientists.

The test results are given in Table X which lists for each of the 19 subjects who took this test, the number tried, the number right and the sigma score for each subtest. Sigma scores are necessarily computed on these distributions.

There are some extremely interesting differences in pattern of scores on this test. Those whose sigma scores are highest on the verbal test include all but one of the botanists, the two anatomists and the physiologist. Those with highest scores on the spatial test include the other botanist, two of the plant geneticists and one animal geneticist. The group highest on the mathematical section includes all of the rest of the geneticists of either variety, and two of the biochemists. The other biochemist and the bacteriologist have spatial and mathematical tests practically equal, with the verbal test lower. Although the groups are small the pattern is a consistent one. It is not surprising that those most adept in nonverbal functions go into genetics and biochemistry, branches in which these skills are put to good use, while these are avoided by the dominantly verbal group. It is interesting, however, that all but two of these men (one botanist ♯8 and one geneticist ♯15) say

that they do the majority of their thinking in terms of visual imagery; only two have dominantly auditory verbal imagery.

TABLE XI

INTERCORRELATIONS OF BIOLOGISTS' SCORES ON THE VERBAL-SPATIAL-MATHEMATICAL TEST

	Number Right		Tried/Right	
	S	M	S	M
V	+.445*	+.129	+.273	—.014
S		+.369		+.385

*For this sample, *P* is .05 for a correlation of .44.

Intercorrelations on the test are given in Table XI. These were run for number tried in each instance to see whether there seemed to be any factor of speed of work, and for the ratio of right to tried as a crude measure of cautiousness. Only one correlation in the table is significant at the 5 per cent level, that between number right on the verbal and spatial tests.

COMPARISONS OF THE THREE TESTS

Although the group is small, this seemed a unique opportunity to make some comparisons of responses on these three tests. An attempt has been made to compare any elements which seem to have meaningfulness. In part, of course, this has been conditioned by the aspects which could be put in quantifiable form, or in terms of frequencies.

It was thought that there might be an underlying factor of productiveness, or general energy output which could be manifested in the amount of work done on each test. For the VSM this could be taken as number tried; for the Rorschach the number of responses, for the TAT the average number of typed lines per story. (This last is fairly crude, and is predicated upon my having gotten down the whole story verbatim. This was almost always the case. In the very few instances in which I could not keep up, the average length will be a little short, but in these terms the differences from actuality are very slight and I think insufficient to distort the results.) These correlations are given in Table XII. None is significant.

TABLE XII

INTERCORRELATIONS OF MEASURES OF PRODUCTIVENESS ON RORSCHACH, TAT, AND VERBAL-SPATIAL-MATHEMATICAL TESTS OF BIOLOGISTS

	Rorschach N Responses	TAT Average Length	V Tried	S Tried	M Tried
Rorschach		+.149	+.364	—.016	+.276
TAT			—.012	+.198	+.313
V Tried				+.210	+.304
S Tried					—.039

TABLE XIII

INTERCORRELATIONS BETWEEN SCORES OF VERBAL-SPATIAL-MATHEMATICAL TEST AND W%, F%, AND F+%, ON THE RORSCHACH FOR BIOLOGISTS

	V Right	S Right	M Right
W%	—.460*	+.001	—.145
F%	+.001	—.105	+.025
F+%	+.085	+.487*	+.140

*P is between .05 and .02.

It seemed of some interest to correlate number right on the VSM tests with some of the Rorschach factors, and for this purpose *W%*, *F%*, and *F+%* were chosen. The results are given in Table XIII. There is, then, a low significant negative correlation between number right on the verbal test and *W%*; and a low positive correlation between *F+%* and number right on the spatial test.

"Originality" on Rorschach and TAT as measured by *O%* on the Rorschach and a measure of *O%* on the TAT secured from the number of unusual stories or additions that are unusual, correlates +.031.

An attempt was made to compare at-

tention to details on the two tests. Subjects whose use of detail on the TAT was particularly noted in the interpretations were compared with the rest as to *D%*, *Dd* and *W%*. The two groups were practically the same on Dd; the more detailed group on TAT was somewhat lower on *D%* and higher on *W%*. The measure on TAT is a very crude one. Nevertheless, it would seem that the perceptual attitude in this respect varies considerably with the material involved.

Munroe's [3] studies of verbal and nonverbal ability as measured by the ACE and Rorschach pattern showed in the students with higher nonverbal ability a significantly greater tendency to use of *F* responses, better form quality and less use of *M* responses. In this group this is not confirmed on *F%*, averages for the subgroups having highest verbal, spatial and mathematical performances being respectively 47.0, 39.5 and 46.1. For *F+%*, however, as might have been expected from the correlations, these averages are 80.3, 94.7 and 76.1 respectively. There is no apparent association with the use of *M* in the subgroups. Numbers of M for the higher verbal group are 1, 1, 1, 1, 3, 3, 9; for the higher spatial, 1, 2, 2, 3, 3, 4; for the higher mathematical, 0, 1, 2, 3, 5, 6. Munroe found subtler differences in the elaboration of responses, and the use of color but these are not quantifiable. Protocols of this group are so different throughout from those given by Sarah Lawrence students that the subtler criteria which made it possible to differentiate the higher verbal and nonverbal groups from their Rorschach protocols cannot be applied to this group. Whether the difference in the protocols are at least in part reflections of the age differences, which is considerable, or of sex differences, cannot be judged.

SUMMARY

Data from three tests, the Rorschach, the Thematic Apperception and a Verbal-spatial-mathematical test given to 20 eminent research biologists have been presented. Wide individual differences appear on all of the tests.

On the Rorschach there is some tendency to increased *W* and also to increased *Dr*. There is some restriction in use of human movement, but no general restriction in that area; there is more general restriction in the color area. Although *F%* is not particularly high generally, there are extremely few compound responses in which *F* is not the primary determinant. Shading shock is generally prevalent from mild to severe in degree. The group is a generally very unaggressive one, and with very little interest in interpersonal relations.

On the TAT there is further confirmation of the generally slow and somewhat restricted psychosexual development, and indications that most of them tend to avoid emotional and interpersonal situations. They are usually unwilling to go beyond the data presented and have a general distaste for the imaginary and a strong preference for concrete realities. Most of them are rather conventional, and have a fairly strong sense of responsibility.

On the VSM, they show all possible patterns of superiority of score on one or the other subtest, but there is a clear separation between those with highest verbal ability and the rest. The first group include all but one of the botanists, and two anatomists and the physiologist. The geneticists and biochemists are without exception better at either the spatial or mathematical test than they are at the verbal.

Except for a correlation of .445 between the verbal and spatial tests, no

significant relationships were demonstrated between the VSM tests. But there was a low, but significant, negative correlation between *W%* on the Rorschach and number right on the verbal test, and a low but significant positive correlation between *F+%* on the Rorschach and number right on the spatial test.

Received January 7, 1949.

REFERENCES

1. KLOPFER, B. AND KELLEY, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1942.
2. MUNROE, RUTH. The inspection technique: A method of rapid evaluation of the Rorschach protocol. *Rorschach Res. Exch.*, 1944, 8, 46-69.
3. MUNROE, RUTH. Rorschach findings on students showing different constellations of subscores on the ACE.. *J. consult. Psychol.*, 1946, 10, 300-315.
4. ROE, ANNE. Analysis of group Rorschachs of biologists. *Rorschach Res. Exch. and J. proj. Tech.*, 1949, 13, 25-43.
5. WYATT, F. The scoring and analysis of the Thematic Apperception Test. *J. Psychol.*, 1947, 24, 319-330.

A PROJECTIVE TEST OF MASCULINITY-FEMININITY[1]

By KATE FRANCK
UNIVERSITY OF CALIFORNIA
AND
EPHRAIM ROSEN
UNIVERSITY OF OREGON

ANYONE attempting to add to the generously sprouting number of tests and "projective" techniques should have a strong conviction that a need exists for even one more test. An increasing number of clinicians believe that linked with other personality problems there is usually found some disturbance of the patient's sex adjustment or concern over his sex role. Existing M-F scales are inadequate because they are highly charged with cultural factors and hence their validity depends upon absence of cultural change [8]. Furthermore, assuming constancy of culturally determined differences between men's and women's interests, activities, information—upon which, for example, the Terman-Miles M-F scale is built—it is readily seen that a woman can be an outstanding seamstress, cook or housekeeper, and also have a disturbing feeling that she is "only" a woman. Indeed, her very outstandingness may be her way of handling the problem. Or a man may be an athlete, may know all about automobiles and fly a plane—and yet be afraid of women. Everyone has known such people, for there are many, who use behavior labeled masculine or feminine by our society to hide their disorientation, often from themselves.

Frequently the Rorschach and TAT are used to fathom such underlying trends. But quantification is difficult, the information available is incidental to other purposes and is often absent in short records.

One of the authors [4] on an earlier occasion investigated the question of masculine versus feminine object choice by men and women. In this study groups of students were confronted with pairs of paintings of sex symbols. They were asked to mark which half of each pair (male symbol or female symbol) attracted them more. It was found that significantly more men than women were attracted by female symbols. The reverse was true for the women.

However, just as homosexual tendencies can spring from many different sources, so attraction to own sex symbols can mean many different things. This first procedure, demanding that the subject merely state his preference between structured stimuli, yields no information about specific meanings. Other means, calling for expression rather than impression on the part of the subject, had to be found for better understanding.

PROCEDURE

In the procedure to be presented here, incomplete drawings, consisting of a few simple lines, are completed by the

[1]The first author constructed and administered the test and formulated the interpretation of data and theoretical basis. Both authors collaborated in establishing criteria, scoring the tests, and writing the article. The second author is responsible for the statistical work.

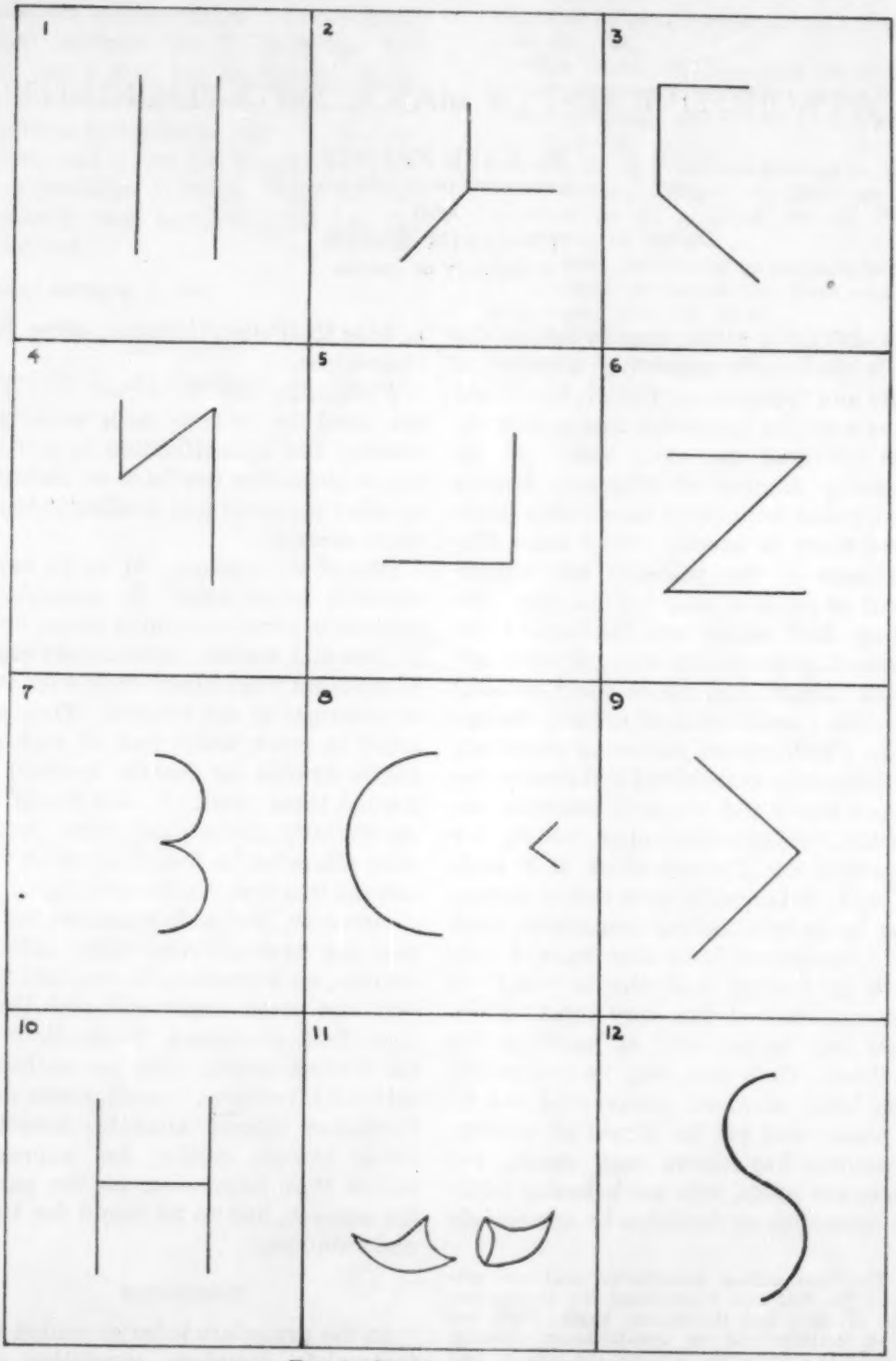

FIG. 1. Sample page of test.

subjects. The stimulus drawings—hereafter referred to as stimuli—extend over approximately one square inch, each located in the center of a cell two and one-quarter inches square. In its present form the test consists of 3 pages, each with 12 stimuli, or a total of 36 items. Instructions are: "These are incomplete drawings. Complete them any way you like. There are no right or wrong answers. Complete the drawings the way you enjoy doing them. You may use as many lines as you please." Deliberately, no effort was made to control the order in which the drawings were completed.

Preliminary investigation of items. Norms were established based on presentation of 60 stimuli to a group of 250 students in an elementary psychology class. Differences between types of completion of drawings by men and by women were examined, formulated as criteria, and tallied for frequency of incidence. These differences, falling into such categories as closure, expansion, internal elaboration, etc., will be discussed in detail later. Thirty-six of the original 60 items yielded statistically significant differences between manner of completion by men and by women. The Chi-square technique was used; P was less than .05 for 28 items; less than .01 for 8 items.

The present test. These 36 items make up the present test. (See Fig. 1 for sample page.) It was given to a new group of 150 female and 150 male students and scored by three scorers according to the criteria established in the preliminary investigation.[2] The completion of each item is scored either masculine: "—"; or feminine "+." The number of "+" scores obtained by an individual is his total score, so that the potential range is from 0 (all items completed in the masculine manner) to 36 (all items completed in the feminine manner). A sample portion of the scoring key is given in Fig. 2. The authors believe that any trained clinician, accustomed to careful scoring and interpretation of projective techniques, can use the scoring manual with little difficulty.

[2]For several items it was found that the original formulation of criteria was somewhat ambiguous, so that reformulation and refinement were necessary.

TABLE I
CORRELATIONS BETWEEN SCORES BY THREE SCORERS

Scorers	Men	Women
K.F. and E.R.	+.90	+.89
K.F. and B.A.	+.86	+.84
E.R. and B.A.	+.85	+.84

RESULTS

Reliability. Correlations between scorers (obtained by the Pearson product-moment r) were computed separately for men and women subjects. The correlations obtained range from +.84 to +.90 (see Table I). The r's for K. F. and E. R., the authors, who together set up the scoring criteria, are slightly higher than between the authors and B. A., who is versed in projective techniques but had no familiarity with the test. The smallness of this difference indicates that little training is required to reach a degree of agreement that is high for a projective technique.

Validity. CR's between men and women were obtained separately for

TABLE II
MEANS, AND DIFFERENCES BETWEEN MEANS FOR MEN AND WOMEN BY THREE SCORERS

	Men		Women		
Scorer	Mean	Sigma	Mean	Sigma	CR
K.F.	14.13	3.43	19.78	3.28	14.12
E.R.	14.80	3.28	19.77	3.35	13.07
B.A.	14.30	3.45	19.27	3.20	13.07

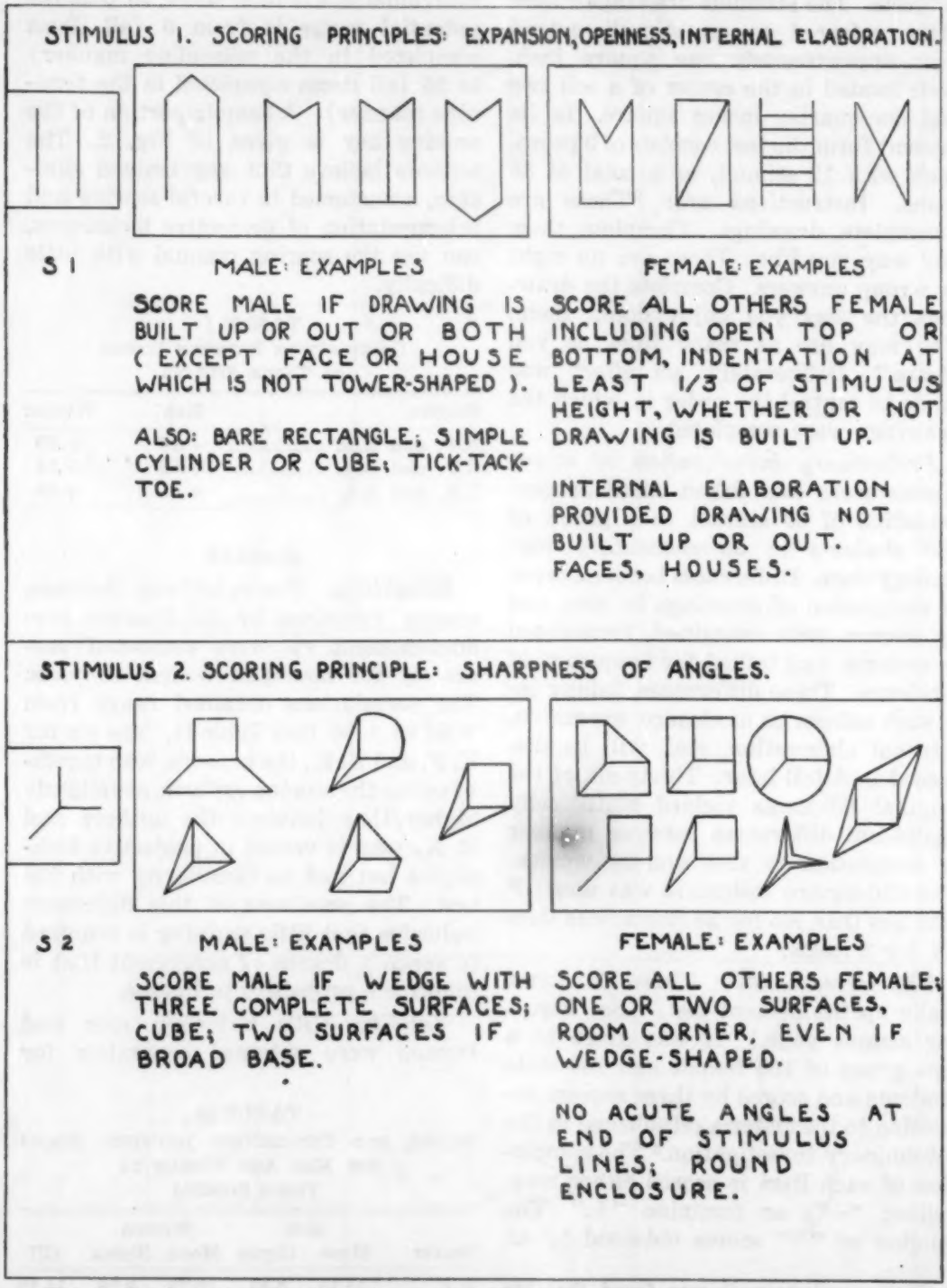
STIMULUS 1. SCORING PRINCIPLES: EXPANSION, OPENNESS, INTERNAL ELABORATION.
S 1
MALE: EXAMPLES
SCORE MALE IF DRAWING IS BUILT UP OR OUT OR BOTH (EXCEPT FACE OR HOUSE WHICH IS NOT TOWER-SHAPED).
ALSO: BARE RECTANGLE; SIMPLE CYLINDER OR CUBE; TICK-TACK-TOE.
FEMALE: EXAMPLES
SCORE ALL OTHERS FEMALE INCLUDING OPEN TOP OR BOTTOM, INDENTATION AT LEAST 1/3 OF STIMULUS HEIGHT, WHETHER OR NOT DRAWING IS BUILT UP.
INTERNAL ELABORATION PROVIDED DRAWING NOT BUILT UP OR OUT. FACES, HOUSES.
STIMULUS 2. SCORING PRINCIPLE: SHARPNESS OF ANGLES.
S 2
MALE: EXAMPLES
SCORE MALE IF WEDGE WITH THREE COMPLETE SURFACES; CUBE; TWO SURFACES IF BROAD BASE.
FEMALE: EXAMPLES
SCORE ALL OTHERS FEMALE; ONE OR TWO SURFACES, ROOM CORNER, EVEN IF WEDGE-SHAPED.
NO ACUTE ANGLES AT END OF STIMULUS LINES; ROUND ENCLOSURE.

FIG. 2. Portion of scoring key. General instructions pertaining to all items are not shown here.

each scorer. They range from 13.07 to 14.12 (see Table II). The range for men was 5 to 24, for women 9 to 28.[3] Using standard scores, an approximate interpretation of an individual score is possible (see Tables III and IV). The division into "extremely masculine," "very masculine," etc., is, of course, crude. However, behavior on this test permits subtler interpretation — to be discussed below—in terms of different kinds of adjustment.

Description of responses. Essentially, then, what are the differences in manner of completion by men and women? Table V summarizes the characteristics of male and female completions. It represents the principles for scoring, stripped of the instructions relating to specific stimuli. (Such specifications are needed because each drawing, starting from a different stimulus, shows the characteristic masculine or feminine manifestations in a slightly different way.) It can be seen that formal criteria outnumber contentual criteria.[4] What few content categories were used partly overlap formal criteria, and may perhaps be discarded eventually.

MEN characteristically *close off* stim-

TABLE III

RAW AND STANDARD SCORES FOR MEN, AND SUGGESTED INTERPRETATION OF POSITION ALONG THE SCALE

Raw Score	Standard Score	Interpretation
0- 3	—4.14 to —3.28	Extremely masculine
4- 6	—3.00 to —2.42	Very masculine
7-10	—2.14 to —1.28	More masculine than average
11-18	—1.00 to +1.00	Average degree of masculinity
19-22	+1.28 to +2.14	Less masculine than average
23-26	+2.42 to +3.28	Feminine tendency
27-36	+3.56 to +6.14	Feminine

TABLE IV

RAW AND STANDARD SCORES FOR WOMEN, AND SUGGESTED INTERPRETATION OF POSITION ALONG THE SCALE

Raw Score	Standard Score	Interpretation
0- 7	—5.56 to —3.56	Masculine
8-11	—3.28 to —2.42	Masculine tendency
12-15	—2.14 to —1.28	Less feminine than average
16-23	—1.00 to +1.00	Average degree of femininity
24-27	+1.28 to +2.14	More feminine than average
28-31	+2.42 to +3.28	Very feminine
32-36	+3.56 to +4.70	Extremely feminine

[3]A group more heterogeneous than our college-student sample would, of course, yield some scores at the extreme ends of the possible range.

TABLE V

CHARACTERISTICS OF COMPLETIONS BY MEN AND WOMEN AS USED IN SCORING CRITERIA

Completed by Men	Stimuli	Completed by Women
Formal categories		
Closing of stimulus	1,2,5,6,10 11,13,15,17, 20,21,25,33, 36	Stimulus left open
Expansion from stimulus outward	1,3,4,5,6 25,26,28	Internal elaboration
Stimulus built up	1,11,13,16, 19,20,22, 23,24	Stimulus built down
Stimulus made sharper	2,9,21,26, 28,29,31	Stimulus made blunter or enclosed
Protruding single line	4,12,16,24, 30	Protruding single lines reenforced or supported
Oneness	8,23,26,27, 35,36	Two-somes
Content categories		
Profiles, nude female figures	7,11,12,14, 15,17,22,23, 26,29,31,35, 36	Frontal faces, figures, stick figures, flowers, fruit, animals
Skyscrapers, towers	1,2,10,14, 18,20,21,2[illegible] 25,27,32,33, 34	Houses, rooms, windows, vases, lampshades, staircases
"Dynamic" objects (those that move under their own power)	19,31,33,34, 14,32,35	"Static" objects (movable by an external force only)

[4]This is due to the authors' belief that formal criteria permit greater generalization and are easier to score than content. There are also theoretical reasons, discussed below, for preferring formal criteria.

ulus areas; they tend to *enlarge, expand* the stimulus, often by means of upward extension. They tend to emphasize *sharp or angular line.* Where a *single line* will do, they make no attempt to double or support it. (Example: See Fig. 3A.) Where there is a choice of unifying or separating the parts of a stimulus, they prefer *unity.* (Example: See Fig. 3B.)

WOMEN leave stimulus areas *open* or elaborate the area *within* the stimulus itself rather than expand (when they do expand it is usually downward). They tend to *blunt* or *enclose* sharp lines or angles, and to *reinforce* protruding single lines as by doubling them. They prefer pairs to unity.

As to the content, *MEN* draw protruding, *closed objects* just as they close off stimulus areas (towers, tools and instruments). They draw *profiles* just as they emphasize angularity.

WOMEN, just as they leave stimulus areas unclosed, draw many *"open" objects* (vases, windows, doors). They draw *frontal faces* rather than profiles, just as they blunt sharp angles.

Independent of formal elements, *MEN* draw "dynamic" objects, i. e., objects that *move under their own power* (automobiles, steamships) or reproduce natural forces (fires, fountains). *WOMEN* draw objects that can be moved, if at all, by some *external force* only (rowboats). They draw all kinds of "containers," as houses, beds, vases. Fruit, flowers, animals, human figures, especially "stick figures" are frequent.

There is a striking similarity between the present findings and the differences observed by E. H. Erikson between boys' and girls' play constructions [2].

DISCUSSION

To understand what prompts masculinity or femininity of response, investigation might proceed from external to central determinants.

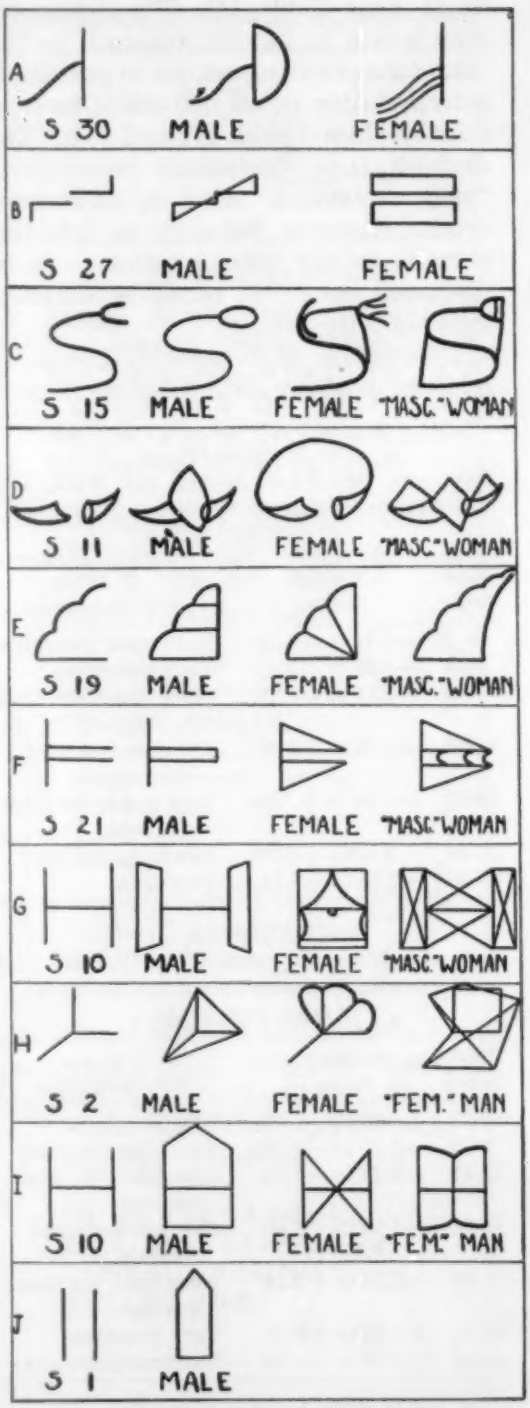

FIG. 3. Some characteristic drawings by different groups.

Cultural factors. Cultural factors, in so far as they prescribe different spheres of interests for men and women, might be held responsible for differences, at least in the content, of the drawings. One might say, for example, that men draw more locomotor objects because they are better acquainted with machines than are women. But they do not draw circular saws or lathes; they draw steamships and automobiles, while women draw rowboats and sailboats. It would be absurd to claim that among college students women are more familiar with sailboats and canoes, and men with ships and cars. Men draw fires and fountains, hardly closer to their spheres of interest than to those of women with their cooking and washing. On the other hand, women draw houses, windows, doors, beds, vases, fruit, flowers. Certainly these constitute part of home and homemaking. However, so do knives, fire pokers, fireplaces, stools and candles—objects women draw rarely. Again, it would be absurd to claim that women have more contact with open objects and men with closed or solid objects.

Cultural factors thus do not seem to be major determinants of the observed responses. (This corresponds, of course, to the authors' intention to construct a scale as free as possible of cultural determinants.)

Freudian symbols. Fruit, flowers and human figures, drawn mainly by women, are equally familiar to both sexes. Most people will agree that they think of flowers as female rather than male. Here the question of symbolic meaning becomes relevant. (Men give women flowers—not vice versa; does this custom owe its origin to the symbolic femininity of flowers?) A good deal has been said for and against the universality of symbols. The authors are inclined to believe that only where universal experience is present can universality of symbols be claimed, as, for example, the use of brightness for happiness, darkness for unhappiness or danger—based on the universal experience of day and night, and the equally universal one of happiness and unhappiness. In view of the universality of sex experience, can sex symbols be expected to be universal? In different environments and cultures, different objects offer themselves as symbols; hence it is dangerous to conclude from the universality of the exprience to be symbolized that choice of symbol is necessarily universal. For these reasons, as well as for simplicity of scoring and interpretation, the writers have emphasized formal rather than symbolic content criteria. When women draw a multitude of open objects and containers it need not be the symbolic aspects of these specific objects that deserve consideration here. It is the common property of being accessible enclosures that has to be scrutinized. The flowers and fruit that women draw have in common their enclosing function in reproduction and their absence of locomotion. Men's preferences for cars, ships, rockets, fires and fountains have in common a dynamic, locomotor property.[5]

Differences in choice of content by the two sexes cannot, then, be understood merely in terms of symbols, but rather as expressions of common properties. In fact, it is possible to reduce symbolic meaning to these abstract properties of form or activity. Therefore, a specific object may be interpreted as expressing masculinity or femininity depending on the manner of drawing the object.

Body structure and function. An attempt will now be made to relate differ-

[5]Of course, the authors do not mean to deny the great psychological difference between the man who chooses ships to represent locomotion and the one who chooses rockets.

ences in the drawings to different body structures and functions of the sexes.

As stated above, men expand from the stimulus outward, tend to build it up, emphasize angles, rely upon the strength of simple lines and prefer oneness. Expansion, upward motion, angular position and oneness are characteristics of the male sex organ, and in a sense, of masculine body build. Dependence upon the strength of the single line may be related to confidence in the organ and in male sexuality. Preference for objects capable of moving under their own power then becomes a logical expression of man's more active sex function.

Women tend to leave the stimulus area open, elaborate within the area and draw round or blunt rather than angular, sharp lines. They prefer twosomes and seem to feel a need to give support to single lines. Accessibility, emphasis on the inside of the body, roundness, symmetry and two-ness are female body characteristics. Flowers, fruit, houses, rooms, vases, bowls are static, "passive," surrounding objects, in the same sense that the womb is receptive and relatively passive. The animals and "stick figures" women draw so often, and that obviously signify children, may then be thought of as an extension of that function.

Relating both formal and content differences to the different structures and functions of the two sexes seems the most satisfactory procedure, since this explanation, of those so far proposed, best accounts for the phenomena observed.

Need and defense. An objection might be raised at this point: the difference in response *between* the two groups is now explained in terms of their different body structures and functions. But what about differences *within* each group? Is it not true that there is infinitely less difference of body structure between the man who obtains a score of 9 (very masculine) and the one who obtains a score of 24 (the most feminine score in the present sample of men) than there is between the man and woman who both obtain a score of 9? Are the man and the woman both earning a score of 9 alike psychologically? Are the man and the woman both earning a score of 24 alike psychologically? Qualitative evaluation will have to furnish evidence to help answer such questions.

Upon inspection, productions by men and women with highly masculine scores show similarity in that both lack the typically female pattern of completion. But here the similarities end. For while the man with a high masculine score draws in the typical masculine manner, the woman with an identical score does not. She is apt to fall into one of two main categories of behavior:

(a) She uses the masculine manner in an exaggerated, absurd way: stimuli are not merely closed off, but are closed off two- and threefold. (Example: See Fig. 3C.) She will consistently and forcibly introduce great angularity. (Example: See Fig. 3D.) Stimuli will be "built up" in a dizzily top-heavy manner. (Example: See Fig. 3E.)

The writers believe that in this pattern behavior of the reaction formation type is manifest: the woman who dares not accept her sex role needs to be not just a man, but needs to make doubly sure that she is "more than a man."

(b) She will, though closing off or expanding, show a muddled indefiniteness of response. She closes off, but not quite. (Example: See Fig 3F.) She expands, but uses internal elaboration simultaneously. (Example: See Fig. 3G.)

Such "confused" drawings may be interpreted as indicating anxiety related to sex role.

Vacillation between clear-cut female responses to some stimuli and fairly clear male ones to others, is a third but rare category. This behavior, the authors believe, may lend itself to interpretation as conflicting male and female identification.

These *a priori* interpretations have been verified in a small number of clinical cases.

However, further research in this direction is in progress.

Conversely, the drawings of men who earn a high female score differ greatly from those of women with high female scores. These men fall into two main categories:

(a) Their drawings are "messy." They use a diffuse multitude of lines where women use simple, better organized internal elaboration. These lines often obscure the stimulus and show no clear response pattern. (Example: See Fig. 3H.) They tend to leave stimulus areas open, but usually have small crowded openings. (Example: See Fig. 3I.)
A possible interpretation is that such behavior indicates an anal pattern, a regression from mature genitality. The anxiety in such patterns is manifest in the crowding of lines.

(b) Their drawings are "playful." They earn their feminine score mainly by drawing houses, rooms, and a great many figures. These differ from women's drawings of the same objects: the men load the houses and rooms with much more detail than women do, their figures tend to be grotesque.
The writers believe that among this group of men will be found many with strong mother identifications, or an attitude of pretended (defensive) castration, as if proclaiming that they were safe because they are women already, a mechanism not unlike that which wards off anxiety in adults by way of humor and in children by way of play. Confirmation from a few clinical cases has been obtained and points the way for future research.

In view of the differences found between drawings by individuals of the same sex, the earlier hypothesis that drawings are simple repetitions of the structure and function of sex organs and body build needs to be modified. Rather, it is suggested that the subjects express something which might be called "body image," in a sense that includes both structure and impulses. This image, which is part of the individual's awareness of his identity and potentialities, corresponds to physiological reality if it is not surrounded and distorted by defensive measures. It follows that the masculinity-femininity test presented here is a measure of degree of acceptance of the individual's sex role, conscious or unconscious.

The authors expect to break the test down for further analysis of disturbances of such acceptance of sex role. This will be done in the way indicated above when "types" of "masculine" women and of "feminine" men were discussed. Prognosis for homosexual tendencies in men—to give an indication for practical use of the test—varies with causative factors. In oversimplified form these factors are often classified as: (a) fear of the opposite sex as dangerous (fear of aggressive mother); (b) fear of the opposite sex as a temptation (fear of being punished by the father for being attracted to the mother); (c) identification with the parent of the opposite sex. Factor (a), it is expected, will manifest itself on the test in initially "normal" drawing, which is then overelaborated, i. e. surrounded by defensive measures; (b), in the authors' opinion, is manifest in the "playful" femininity discussed above under Freudian symbols; (c) will produce behavior falling between the drawings of group (b) and drawings by women. (Prognosis, of course, is best for (a), worst for (c).)

THEORETICAL CONSIDERATIONS

Is this test, a "projective technique," actually based on projection? The term projection is used by Freud [5], Fenichel [3], Knight [6] and others for behavior in which the individual assigns to the external world processes and images that really occur within himself. Given opportunity on the one hand, and tension on the other, anything at all may be projected: images of goal objects as well as impulses toward those objects, fears as well as defenses.

The different projective techniques probably tend to elicit more of one or the other, goal objects or impulses: the Rorschach, for example, invites visual imagery, and hence projection of goal

objects more than of impulses toward action. Recognizing projection of goal object as different from projection of impulse is most important: the male response, for example, on S1 is often "built up" (see Fig. 3J); the result looks much like a phallic form. If this be interpreted as projection of a goal *object* one would have to conclude either homosexual wishes or a severe castration delusion. Such a conclusion, if correct, would mean that a large percentage of the male population is severely disturbed, or expressed another way, that such disturbance is the norm. However, this test, demanding the active motor involvement of drawing, appears to elicit "projection" of motor activity. If then, drawings like S1, as shown in Fig. 3J, are considered as expressing upward motion—and similar manifestations are considered in the same manner—then what has been "projected" was simply an impulse toward sex activity, in some cases with the anxiety and the defenses connected with it.

It should be remembered at this point that projection is essentially a defense mechanism: that which is ego-alien or disturbing is gotten rid of by externalization. The usefulness of projective techniques is based on this principle of bringing to the fore, above all else, that which is disturbing the individual. Yet, projection as a defense is spontaneous, while projective tests are given alike to those who use projection as a major defense and to those for whom it is of secondary importance. Not only is the need to project present to very varied degrees, but also the "projected" material is ego-alien or disturbing only in part. For example, it can hardly be assumed that the sex functions expressed in the drawings of the majority of subjects are felt as disturbing or foreign. It seems more reasonable to assume that some tension is attached to sexual needs in general, which is partly relieved by being expressed. Such partial relief of a basic need, in a constructive manner, may be considered self-expression, or more precisely self-representation—or self-recreation, in Kris' [7] sense of the word—rather than projection. The writers suggest that the masculinity-femininity scale here presented is, incidentally, a scale representing a continuum from projection to constructive self-representation.

To an unknown extent this enters, of course, into most productions on projective tests, although it usually eludes quantification. The drawing completion test, by making self-expression more explicit, holds a position somewhere between a projective technique and creative art, with the former showing maximum projective behavior, the latter maximum sublimating behavior.

Received January 3, 1949.

References

1. Erikson, E. H. Clinical studies in childhood play. In R. G. Barker, J. S. Kounin and H. F. Wright, (Eds.) *Child behavior and development.* New York: McGraw-Hill, 1943. Pp. 411-428.
2. Erikson, E. H. Further explorations in play construction: three spatial variables in relation to sex and anxiety. *Psychol. Bull.* 1941, 38, 748 (Abstract).
3. Fenichel, O. *The psychoanalytic theory of neurosis.* New York: Norton, 1945.
4. Franck, Kate. Preference for sex symbols and their personality correlates, *Genet. Psychol. Monogr.* 1946, 33, 73-123.
5. Freud, S. *Collected papers.* London: Hogarth Press, 1946. Vol. II.
6. Knight, R. P. Introspection, projection and identification. *Psychoanal. Quart.* 1940, 9, 334-341.
7. Kris, E. Approaches to art. In S. Lorand (Ed.) *Psychoanalysis today.* New York: International Universities Press, 1944. Pp. 354-370.
8. Terman, L. M. and Miles, Catharine C. *Sex and personality.* New York: McGraw-Hill, 1936.

STATISTICAL TESTS OF CERTAIN RORSCHACH ASSUMPTIONS: ANALYSES OF DISCRETE RESPONSES

By J. R. WITTENBORN
YALE UNIVERSITY

THE customary procedures employed in the evaluation of a subject's responses to the Rorschach ink blots involve a variety of assumptions. Some of the assumptions are explicitly stated in the Rorschach literature and others are merely implied by common practices. Although the validity of the Rorschach technique for personality appraisal is indicated by clinical findings and certain types of studies, the assumptions employed in the Rorschach psychodiagnosis have not been subjected to direct validation. There are considerations which make particularly important the validation of the assumptions per se.

1. Since the assumptions have not been tested, the possibility exists that some of them are invalid. If some of the assumptions cannot be validated, aspects of scoring or interpretation based on the questionable assumptions should be employed with reservation or perhaps eliminated from standard procedures.

2. Many of the assumptions employed in Rorschach procedures cannot be found among the established facts and theories of American academic psychology. If the Rorschach assumptions are valid, their relevance to the psychology of personality organization and to the psychology of perception could be of extraordinary significance.

ANALYSIS A. AN EXAMINATION OF THE FUNCTIONAL SIMILARITY OF RESPONSES WHICH HAVE COMMON DETERMINANTS

In scoring a Rorschach protocol, it is customary to classify each response with respect to several classes of factors, such as content, portion of ink blots interpreted, and ink blot characteristics determining the response. There are standard categories within each of the classes of factors. When a response is scored the content category in which it falls, as well as its location category and its determinant category, is designated by the scorer. As a second step in scoring, the total number of responses falling in each category is ascertained. Regardless of the precise manner in which the categories are used, the device of categorizing the responses and then finding the total number of responses in each category carries two implications:

I. That all of the responses falling in a given category are similar in some behavioral respect.

II. That the psychological significance of responses falling in a given category is different in some respect from responses placed in other categories.

These assumptions are implicit for example in the distinction that is made between Color and Human Movement responses. The importance of the distinction between Human Movement and Color responses was first emphasized by Rorschach, and the observance of this distinction has remained a basic feature of Rorschach psychodiagnosis. Because of the basic significance ascribed to these two types of responses, it

was decided to employ this distinction in testing the foregoing assumptions of intracategory similarity and intercategory difference.

It would be erroneous to suppose that any absolute or relative quantity of color or movement responses is unalterably and invariably associated with a particular total personality appraisal. Movement and color have an invariable abstract meaning for the personality, but for total personality appraisals this is added to, detracted from, or modified by other characteristics of the response and other aspects of the protocol. This is analogous to clinical practice wherein the psychologist's evaluation may be based on the Rorschach indications, other test data, and data from the case history; the relative contribution ascribed to the Rorschach in the total personality evaluation is determined to some degree by the other findings, but this does not render a Rorschach finding meaningless. By the same token the qualifying role of such factors as form level and content do not make the concept of a movement response meaningless nor make inconsequential an assumption that movement responses include a constant, functionally similar element.

THE PLAN OF EXPERIMENT

If Assumption I is a true proposition, it follows that several measurements of the tendency to perceive Human Movement would be positively related. If several measures of a given response category were not positively interrelated, it would be difficult to justify an assumption that differences in number of responses of that category expressed differences in a personality attribute. It also follows from Assumption II that several measures of a given response category would be more highly interrelated with each other than they would be related with measures of some other category.

In order to employ these deductions in an experimental test of the assumptions, it is necessary to devise several measures of the tendency to give responses falling in the respective categories. For example, it is possible to regard any response falling within a given category as evidence of a tendency to give responses in that category, and it is equally possible to regard failure to give a particular response falling in a given category as evidence of a tendency not to give responses falling in the category. Employing this approach, it is possible to secure a large number of plus or minus scores for any response category. If making or not making a particular response which belongs to a given category is to be a measure for a tendency to make respectively many or few responses which fall in that category, it is important that the particular response be well specified. This poses an important methodological problem.

One of the outstanding characteristics of Rorschach responses is their highly individual quality, their extreme variability from person to person. This makes conventional statistical treatment of Rorschach data awkward because of the difficulty in securing a sufficiently large number of roughly identical responses. For example, out of a group of one hundred normal subjects many, if not all, would report Human Movement responses, but relatively few would have more than two or at the most three identical responses in common.

In the present study an attempt has been made to reduce this difficulty by use of the Harrower-Erickson check list [6]. This check list comprises a sheet of paper upon which are printed three groups of ten possible responses for each of the ten cards. There are three hundred possible responses in all.

The technique of administration provides that the ten cards be projected on a screen in the usual serial order. The check list is designed so that the subject may check or underline responses which indicate what he has seen on the card. Although this technique of administration and manner of responding to the cards is obviously different from the standard technique for administering the Rorschach test, this device provides a uniform, common response list for all subjects and thus makes possible a tabular treatment of responses to ink blots. The limitation of this procedure as a method for personality appraisal [3, 5, 7] is not necessarily disadvantageous to the present experiment which is an examination of the consequences of certain restricted assumptions and is not otherwise concerned with the problem of the validity of the Rorschach method in general or the validity of other Rorschach assumptions.

The following experimental hypotheses were tested by a statistical analysis of data provided by the Harrower-Erickson check list:

A. The tendency for controlled Color responses (FC and CF) to be associated with other controlled Color responses (FC and CF) is greater than the tendency for such Color responses to be associated with Human Movement responses.

B. The tendency for Human Movement (M) responses to be associated with other Human Movement (M) responses is greater than the tendency for such responses to be associated with controlled Color responses (FC and CF) .

RESULTS

The subjects for the investigation were 247 Yale University freshmen, all the members of a small entering class. The check list was administered as a part of a guidance battery. The responses selected for analysis had to meet several requirements. The responses had to be sufficiently common to afford a basis for statistical tests. The permissible range of response frequency was arbitrarily set from 28 to 220. The exact range was influenced by efforts to meet certain other requirements, e.g., an effort was made to include responses from all cards; different responses to the same portion of the card which could

TABLE I

CHI-SQUARE TEST OF INDEPENDENCE BETWEEN RORSCHACH RESPONSES

6.63=1% level 3.84 = 5% level 2.71 = 10% level

No.	Response	Card No.	Resp. Group	1	2	3	4	5	6	7	8	9	10	11	12
1	Two Clowns	II	a												
2	Two Men Pulling Something Apart	III	a	2.1											
3	Man Seen From Below	IV	c	3.6	1.3										
4	A Fan Dancer	V	c	3.2	3.3	14.7									
5	Two women Talking	VII	a	13.4	.2	1.4	1.3								
6	Two witches	IX	a	3.3	.9	5.1	2.3	5.5							
7	Two People	X	a	.3	.9	14.7	5.9	4.6	10.4						
8	A Red Bow Tie	III	a	.1	.5	3.3	12.4	1.4	14.3	3.6					
9	An Emblem	VIII	a	.07	.1	2.1	1.9	.3	1.6	3.2	2.1				
10	Colored Map of California	X	a	.1	.0	2.4	5.4	.1	1.3	1.0	9.6	3.2			
11	Red and Black Ink	II	a	3.3	.3	.0	.5	16.1	1.7	.2	.0	.4	1.5		
12	Fire and Ice	VIII	b	4.4	.0	3.5	1.5	1.0	9.7	.3	6.4	.2	9.2	.2	
13	Forest Fire	IX	b	2.7	.2	.2	3.0	.7	.3	.3	.5	1.1	1.3	1.3	3.3

therefore be construed as mutually exclusive were (with two noted exceptions) avoided. The responses finally selected are listed in Table I.

The color responses comprise two groups which differ in the degree to which they are influenced by the form characteristics of the card. Responses numbers 8, 9 and 10 are strongly influenced by form and would be scored FC, whereas responses 11, 12 and 13 are less influenced by form and would be scored CF by most workers. The color responses are not strictly comparable with respect to the location of the response; for example, responses 8 and 10 clearly involve large details of the cards, whereas with the exception of response 9, which is probably based on the whole card, the others are less definite in their exact location. For the most part the Human Movement responses are based on most if not all of the total area of the respective ink blots. Responses 1, 6 and 7 are based on colored cards and they may involve color to some degree. This possibility is greatest for response 1; clowns (or equivalent movement responses for card II) are commonly given a specific identification which employes the red color. It is possible to control these complicating and extraneous factors by restricting comparisons to certain of the selected responses. For example, Human Movement responses 3, 4 and 7 comprise practically all of the ink blot and cannot involve color because the blots are achromatic. Similarly, responses 11, 12 and 13 are homogeneous in that they tend to involve most of the card and would usually be scored CF.

In order to test the experimental hypotheses, each response was correlated with every other one. Two different statistical devices were employed to examine the relationships. The χ^2 test of independence was first used. The size of χ^2 is not determined by the degree of relationship, but since the frequency with which χ^2 of any given size may occur by chance is known, χ^2 may be used as a test for the significance of the correlation, i.e., the lack of independence between two responses may be evaluated by means of the χ^2 test. The result of this analysis is shown in Table I.

In order to get some appreciation of the strength of the significant relationships revealed by the χ^2 test, tetrachoric correlations between the responses were computed. The correlations were determined by means of a chart [4]. Since the tetrachoric correlation is an appropriate statistic only when the dichotomized variables are normally distributed, its use for the present purpose may be questioned; whether the strength of a tendency to give a Rorschach response is normally distributed or not is a speculative question. At any rate the indications provided by the tetrachoric correlations are in agrement with those provided by the χ^2 test. The tetachoric correlations between all possible pairs of responses are given in Table II.

DISCUSSION

Two conclusions seem equally appropriate from an examination of either Table I or Table II:

1. Among the intercorrelations there is no obvious pattern which conforms with a pattern predictable from either Hypotheses I or II.
2. The number of significant correlations exceeds the number that would be expected by pure chance.

Although it would be unwarranted on the basis of the present data to make any conclusive statements concerning the general status of Assumptions I and II, the hypotheses generated by them for this experiment are not clearly verified. If the Rorschach were analagous to the typical mental test wherein the

score is based on many responses of a given kind and evaluation of the individual is based on differences in many score units, a high degree of interitem consistency would not be necessary. But such is not the case, and in view of the considerable psychological significance ascribed to small variations in number of Human Movement or Color responses, highly consistent inter-response evidence for the hypotheses may be expected. The scant evidence for the assumptions is a challenge to their status and strongly indicates the need for further investigation.

The fact that among the responses there are numerous relatively large and significant interrelationships which were not predicted by the hypotheses indicates need for a broader examination of the patterning of Rorschach responses. It is possible that the interrelationships among Rorschach responses have a well defined and meaningful pattern which if discovered would contribute to our understanding and use of the Rorschach technique.

Although from a casual examination of Tables I and II it appears that the evidence which the present data provide for the assumptions is unequal to the important manner in which they are employed in practice, any evidence of a trend toward correspondence between the requirements of the hypotheses and the nature of the data should not be overlooked.

Accordingly, Table III was prepared to summarize the findings and from

TABLE III

DISTRIBUTION OF CHI-SQUARES AND TETRACHORIC CORRELATION COEFFICIENTS

(A Summary of Tables I and II)

χ^2	% of the 42 possible Comparisons of Human Movement with Human Movement	% of the 42 possible Comparisons of Human Movement with Color*	% of the 30 possible Comparisons of Color with Color
Less than 2.71	42	69	67
2.71 and above (10% level)	58	21	33
3.84 and above (5% level)	38	19	27
6.635 and above (1% level)	19	15	13
Tetrachoric r			
Less than .20	33	62	60
.20 or above	67	38	40
.30 or above	24	21	13
.40 or above	10	05	00

*Two of the significant χ^2 are for negative relationships (see Tables I and II.)

TABLE II

TETRACHORIC CORRELATIONS BETWEEN RORSCHACH RESPONSES

No.	Response	Card No.	Resp. Group	1	2	3	4	5	6	7	8	9	10	11	12
1	Two Clowns	II	a												
2	Two Men Pulling Something Apart	III	a	.11											
3	Man Seen From Below	IV	c	.20	.20										
4	A Fan Dancer	V	c	.21	.16	.42									
5	Two women Talking	VII	a	.32	—.01	.16	.17								
6	Two Witches	IX	a	.20	.20	.22	.23	.26							
7	Two People	X	a	.08	—.10	.40	.27	.30	.33						
8	A Red Bow Tie	III	a	.00	—.14	.30	.27	.15	.43	.23					
9	An Emblem	VIII	a	.08	—.07	.13	.14	.02	.15	.30	.15				
10	Colored Map of California	X	a	.00	—.05	.14	.24	.03	.20	.10	.20	.22			
11	Red and Black Ink	II	a	—.38	.01	—.10	—.10	—.50	.07	.06	—.01	.02	.13		
12	Fire and Ice	VIII	b	.33	.05	.34	.16	.20	.40	.33	.31	.07	.31	.09	
13	Forest Fire	IX	b	.20	—.05	—.07	.44	.20	.07	.22	—.10	.21	.02	—.15	.20

this summary several statements may be offered:

1. There is no consistent functional similarity among the Human Movement responses or among the Color responses, i.e., most of the movement responses are not significantly interrelated and similarly most of the Color responses are not significantly interrelated.

2. The pattern among the most significant relationships is in a direction predictable from the hypotheses.

3. The hypotheses are not clearly verified and the assumptions may be false.

As a result of the analysis it is concluded that in these data the tendencies required by hypotheses A and B are negligible. A variety of considerations may be adduced to account for negligible evidence for experimental hypotheses. Among them the following may be included:

1. The assumptions may be at fault.

a) The assumptions may be false (The present findings merely challenge the assumptions).

b) The assumptions may describe a very slight trend. (The present evidence does suggest that the trend is much less marked than might be inferred from Rorschach practices).

c) The statement of the assumptions may be faulty. (One of the purposes of research is to arrive at correctly stated assumptions, and the refutation of hypotheses is the basis for reformulation of assumptions).

2. The experimental hypotheses may not be the logical consequences of the assumptions.

a) The hypotheses may involve assumptions other than those under examination. (The present experiment does involve assumptions regarding the meaning and scalability of a response).

b) The hypotheses may be incorrectly deduced, e.g., so as to be in conflict with other possible deductions.

3. The experimental interpretation may be faulty.

a) The data may be partially or wholly irrelevant to the hypotheses. (The present data may be inappropriate to a general examination of Rorschach assumptions, but they are considered appropriate to the present assumptions).

b) The data may be influenced by constant, uncontrolled effects (either the group situation or the list of possible responses may have an untoward effect on the individual's responses).

c) Relative to the magnitude of the trend to be examined, errors of measurement and other variable extraneous factors may be large and thus obscure the findings. (This possibility is not eliminated in the present investigation, e.g., other aspects of the personality not expressed through movement or color may vary from response to response and from individual to individual).

ANALYSIS B. AN EXAMINATION OF THE FUNCTIONAL SIMILARITY OF A LARGE NUMBER OF CHECK LIST RESPONSES

The major conclusions of Analysis A may be stated as follows:

1. Among the intercorrelations there is no obvious pattern which conforms with a pattern predictable from either Hypothesis I or Hypothesis II.

2. The number of significant correlations exceeds the number that would be expected by pure chance.

These conclusions, particularly the latter one, provide the point of depar-

ture for a second analysis. Since there is a greater number of significant intercorrelations among the Rorschach responses than chance alone would provide, it is possible that the pattern of the intercorrelations could be anticipated to some degree from a consideration of certain beliefs and practices concerning the Rorschach. As indicated in the discussion of the results of Analysis A, the consistent behavioral significance among Human Movement or among Color responses may be relatively small compared with the variations among the responses due to other inconsistent factors, e.g., a group of responses selected because they all involve color may differ greatly in other respects. For example, some of the responses will involve the whole card whereas others will be based on a small detail, and the contents of some of the responses may be man-made objects while others may be plants, etc. It may be reasoned, therefore, that the intercorrelations among responses which have more than one scoring category in common will be higher than the intercorrelations among responses which have but one or no common scoring category. These considerations lead to a third hypothesis for testing the provisions of Assumptions I and II.

Hypothesis C: The degree to which Rorschach responses will be intercorrelated is a function of the number of scoring categories the responses have in common.

Since it is obviously infeasible to study the interrelationships among all of the 300 responses possible in the Harrower-Erickson check list, some selection was necessary. In selecting the responses for Analysis II several factors were considered:

1. The responses must be sufficiently common to make statistical analysis feasible.

2. The manner in which they are scored must be unambiguous.

3. They must be pertinent to common Rorschach practices and assumptions.

4. Whenever possible, the responses to be studied should be selected from the first group of 10 possible responses for each card or at least the response of a given type selected for a card was the first one listed for each card.

5. Within these restrictions the responses selected should represent a variety, and not be limited to one or two types.

As a result of these considerations, twenty-seven check list responses were selected from the records of a small Yale Freshman Class of 247 and intercorrelated by the tetrachoric method.

As a second step in the analysis, a location, determinant, and content scoring was made for each response. The scoring for each response was then compared with the scoring for every other response and for each response four lists were prepared:[1]

a) The responses which agreed in three respects (had location, determinant, and content in common).

b) The responses which agreed in two respects.

c) The responses which agreed in one respect.

d) The responses which agreed in no respect (had no common scoring category).

On the basis of these four lists, four distributions of correlation coefficients were made:

a) The three-degree-of-correspondence distribution comprised all the correlations between each response and every other response which had identical scoring categories. (Table IV).

b) The two-degree-of-correspondence distribution comprised all the correlations between each response and every other response wherein two scoring categories were shared (Table IV).

c) The one-degree-of-correspondence distribution comprised all the correlations be-

[1]The correlations between responses based on a common card were not included in the analysis.

TABLE IV

FREQUENCY DISTRIBUTIONS OF CORRELATION COEFFICIENTS BETWEEN RESPONSES WITH VARYING NUMBERS OF ASPECTS IN COMMON

Interval of r	Frequency Distributions of Correlation Coefficients Between Responses With: Three Aspects Same	Two Aspects Same	One Aspect Same	No Aspect Same	Total Distr.
+ .4 - + .5	2	4	14	4	24
+ .3 - + .4	12	6	18	18	54
+ .2 - + .3	20	18	70	22	130
+ .1 - + .2	28	10	74	50	162
.0 - + .1	14	12	92	36	154
— .1 - .0	8	4	42	18	72
— .2 - — .1	—	2	28	10	40
— .3 - — .2	—	—	—	—	—
— .4 - — .3	—	—	—	—	—
— .5 - — .4	—	—	2	—	2
— .6 - — .5	—	—	—	—	—
— .7 - — .6	—	—	—	—	—
— .8 - — .7	—	—	—	—	—
— .9 - — .8	—	—	2	—	2
	N = 84	N = .56	N = 342	N = 158	N = 640
	M = + .176	M = + .184	M = + .108	M = + .130	M = + .129
	S = .123	S = .147	S = .168	S = .143	S = .176

tween each response and every other response having but one common scoring category (Table IV).

d) The zero-degree-of-correspondence distribution comprised all of the correlations between each response and every other response having no common scoring category (Table IV).

Hypothesis C is tested by means of the χ^2 test shown in Table V. This table shows that there is a significant tendency for the number of response categories common to pairs of Rorschach responses to be positively related with the size of the correlations between the paired responses. The trend does not appear to be a continuous one, however, and as a means of examining it more minutely, each degree of correspondence was compared with every other one. Responses which have two or three common scoring categories have probable higher intercorrelations than responses which have one or less common scoring categories. There is no significant difference between the size of the correlations between responses which have two common categories and the responses which have three common categories: nor is the difference between the size of the correlations between responses having one common category and the size of the correlations between responses having no common categories significant. These findings are taken as evi-

TABLE V

CHI-SQUARE BETWEEN VALUES OF CORRELATION COEFFICIENTS AND NUMBER OF RESPONSE-ASPECTS IN COMMON

Value of r	Number of Respects in which Responses are Same: 3 Frequency	2 Frequency	1 Frequency	0 Frequency	Total
.20 plus	34 (28)	28 (19)	102 (114)	44 (52)	208
.0 - .2	42 (41)	22 (27)	166 (167)	86 (77)	316
less than .0	8 (15)	6 (10)	74 (61)	28 (28)	116
total	84	56	342	158	640

n = 6 $\quad \chi^2 = 18.932$ $\quad P = .01$

dence for Hypothesis C.

The relatively small differences between the size of the correlations in each of the four distributions (Table IV) and the variability of each of the distributions is meagre support for the practice of categorizing the small number of responses provided by the usual Rorschach protocol. The large number of zero or negative correlations between responses which have two or even three scoring categories suggests that the trend provided by Assumption I is indeed weak. The number of correlations above .30 between responses which have one or even no common category suggests that the distinction between responses having different scoring categories (Assumption II) could be of little or no practical value in appraising individuals on the basis of these data.

DISCUSSION

As a result of the Analysis B two conclusions are permissible:

1. There is a statistically significant tendency for Rorschach responses which have two or three scoring categories in common to be more highly intercorrelated than responses which have but one or no common scoring categories. Thus Hypothesis C is verified.

2. The tendency for the correlations between responses to vary directly with the number of common scoring categories is so negligible as to be associated with a large number of gross exceptions. The number and size of the exceptional correlations is such as to challenge the propriety of the practice of scoring the Rorschach check list responses on the basis of location, determinant, and content.

The present study has been concerned with the functional similarity of responses which are scored in a highly similar if not identical manner. The relevance of this analysis is based on the presumption that responses which receive an identical designation on the basis of a complex categorization should have a strong, consistent, demonstrable, functional similarity if the categorization has implications for making distinctions between people. It is similarly argued that responses which are placed in dissimilar scoring categories should have a relatively low degree of functional similarity. The result of the present analysis is not in conformance with these expectations, and the value of the practice of scoring Rorschach responses on the basis of location, determinant, and content is hereby questioned. It is important to note that the data employed in the present study were secured by the use of a check list. The use of a check list was necessary to a study such as the present one, and it is to be emphasized that the inability of the examiner to make qualitative judgements of the examinee's behavior, to make an inquiry, test limits, and conduct a searching interview with the client regarding his background and present difficulties (such interviews appear to be the rule, rather than the exception with Rorschach examiners) is beside the point of the present study. There are, however, two possible respects in which the check list technique would elicit atypical data:

1. Actually seeing the possible responses may in some cases suggest a response which would not be made in the standard situation.

2. The presence of others in the testing situation may result in responses which might otherwise not appear.

These possibilities impose a limitation in the degree to which the present results may be generalized to the standard individual Rorschach procedure.

SUMMARY OF ANALYSES A AND B

Analyses A and B comprise an examination of the responses of two hundred and forty-seven students to a Rorschach check list. The analyses were designed to determine empirically the status of certain hypotheses generated by two assumptions which are relevant to practices in Rorschach scoring and interpretation. The assumptions from which the hypotheses were deduced are:

Assumption I, All the responses falling in a given Rorschach scoring category are similar in some behavioral respects.

Assumption II, The psychological significance of responses falling in a given scoring category is different from that of responses in other categories.

Analysis A was designed to examine the status of Hypotheses A and B. The analysis of the data offers negligible support for the validity of these hypotheses:

Hypothesis A, The tendency for controlled Color responses (FC and CF) to be associated with other controlled Color responses (FC and CF) is greater than the tendency for such Color responses to be associated with Human Movement responses.

Hypothesis B, The tendency for Human Movement (M) responses to be associated with other Human Movement (M) responses is greater than the tendency for such responses to be associated with controlled Color responses (FC and CF).

Analysis B provided a test for Hypothesis C:

Hypothesis C, The degree to which Rorschach responses will be intercorrelated is a function of the number of scoring categories the responses have in common.

The evidence for Hypothesis C met a criterion for a high order of statistical significance. Nevertheless, the trend was slight and marked by numerous important exceptions (see Table IV). Although the data support Hypothesis C they may be interpreted as comprising a challenge to the clinical value of current Rorschach scoring procedures rather than as an indication that the procedures are economical and highly efficacious in attempts at appraising the individual.

The data employed in the present study are check list data. Had they afforded strong support for the hypotheses, the implications would have been readily generalized to include responses elicited by individual methods of Rorschach examination. Since the data failed to offer strong support for the hypotheses, however, the findings are regarded with caution. Analyses must be designed which can be conducted with material from individually conducted protocols. It is possible that exposing a list of alternative responses to a subject biases him (particularly if he is suggestible), and may detract from trends which could be clearly demonstrated in material from individually administered Rorschach tests. It seems unlikely, however, that well-defined patterns of response could be revealed in one testing situation and altogether obscured in another. It is suggested, therefore, that the present data may be taken as evidence that the provisions of Assumptions I and II will not be sufficiently marked in individually administered protocols to warrant the confident use which is constantly made of them.

The scant support which the Hypotheses A, B, and C receive is taken as indication that the usual abstract scoring procedures are of no value in attempts to appraise the behavioral significance of Rorschach responses elicited by check list procedures.

Received November 26, 1948.

REFERENCES

1. BALINSKY, B. The multiple choice group Rorschach test as a means of screening applicants for jobs. *J. Psychol.*, 1945, 19, 203-208.
2. BECK, S. J. *Rorschach's test.* New York: Grune and Stratton, 1945. Vols. I and II.
3. CHALLMAN, R. C. The validity of the Harrower-Erickson Multiple Choice test as a screening device. *J. Psychol.*, 1945, 20, 41-48.
4. CHESIRE, L., SAFFIR, M. AND THURSTONE, L. L. *Computing diagrams for the tetrachoric correlation coefficient.* Chicago: Univ. of Chicago Bookstore, 1938.
5. ENGLE, T. L. The use of the Harrower-Erickson Multiple Choice test in differentiating between well adjusted and maladjusted high school pupils. *J. educ. Psychol.*, 1946, 37, 550-556.
6. HARROWER, M. R.-ERICKSON AND STEINER, M. E. *Large scale Rorschach techniques.* Springfield, Ill.: Charles C. Thomas, 1945.
7. JENSEN, M. B. AND ROTTER, J. B. The validity of the multiple choice Rorschach test in Officer Candidate selection. *Psychol. Bull.*, 1945, 42, 182-185.
8. KLOPFER, B. AND KELLEY, D. M. *The Rorschach technique.* Yonkers, N. Y.: World Book Co., 1942.
9. LAWSHE, C. H., JR. AND FORSTER, MAX H. Studies in projective techniques: I. The reliability of a multiple choice group Rorschach test. *J. appl. Psychol.*, 1947, 31, 199-211.
10. RAPAPORT, D. *Diagnostic psychological testing.* Chicago: Year Book Pub., 1945. Vols. I and II.
11. RORSCHACH, H. *Psychodiagnostics.* Berne: Hans Huber, 1942.
12. WINFIELD, M. C. The use of the Harrower-Erickson Multiple Choice Rorschach test with a selected group of women in military service. *J. appl. Psychol.*, 1946, 30, 481-487.

VALIDITY OF THE RORSCHACH 8-9-10 PER CENT AS AN INDICATOR OF RESPONSIVENESS TO COLOR

By BERT R. SAPPENFIELD and SAMUEL L. BUKER
MONTANA STATE UNIVERSITY

MANY psychologists have been dubious about the value of the Rorschach test, perhaps primarily because its validity has been accepted on faith by a number of its enthusiastic proponents prior to a thorough-going experimental analysis.

That this point of view is well grounded is exemplified by the use of the 8-9-10 per cent as an indicator of *Erlebnistyp*. Klopfer [1, pp. 254-256], for example, states:

> The percentage of responses to cards VIII, IX, and X, seems to indicate a responsiveness to stimuli from without which is even less under the conscious control of the subject than the use of action and color elements. . . . The cause of underproduction can be associated with color effects in two ways; either the color merely has no particularly stimulating effects, or the color actually has a disturbing effect on the subject.

Klopfer establishes 30 per cent as the point below which underproduction is indicated and 40 per cent to 50 per cent as the zone above which overproduction in indicated. It is clear that Klopfer makes an unverified assumption that productivity on the last three cards is a function of responsiveness to color. The possibility exists that productivity on these cards may be attributable to some other factor than color.

The purpose of this report is to describe an experiment designed to test the validity of Klopfer's assumption. Our hypothesis was that, if productivity on blots VIII, IX, and X is truly a function of responsiveness to color, then productivity on the last three cards of a completely achromatic series of Rorschach blots should not differ from productivity on the other seven blots, since in this case there would be no color differences among the blots. That is, the 8-9-10 per cent would be expected to average about 30 per cent for a group and to show a minimum of individual variation from this average value. The reasoning of Klopfer and other Rorschach workers involves the assumption that color differences constitute the primary factor accounting for variability in 8-9-10 per cent from the expected value of 30 per cent. If a similar degree of variability in 8-9-10 per cent should occur when all blots are achromatic, then this variability would require interpretation in terms other than differences in color.

PROCEDURE

Because the design of the experiment depended upon presentation of an achromatic series of Rorschach blots it was necessary to produce such a set, since none is commercially available. In the commercially produced Harrower-Erickson Kodaslides, blots I, IV, V, VI, and VII are in achromatic form. Achromatic slides for cards II, III, VIII, IX, and X were reproduced locally by photographic means. Cards II and III were reproduced on Super XX at f. 1/32 and cards VIII, IX, and X on XF Pan at f. 1/22, and mounted on Kodachrome binders. It was necessary to use the

contrast XF Pan for cards VIII, IX, and X because the brightness values of the figures on these cards are very close to the brightness values of the white backgrounds, and normal film will not record the same contrast that is apparent when hue is present.[1]

Each subject was shown both the Harrower-Erickson series of Kodaslides and the achromatic series on separate days at an interval of one week. Approximately half of the subjects responded to the Harrower-Erickson series first and the rest to the achromatic series first. Each slide was exposed for 2.5 minutes in an upright position in a darkened classroom. Vision was good from all seating positions in the room. Subjects made written responses on a special form. This form provided a separate page for each slide, each page bearing a mimeographed reproduction of the blot. After exposing all ten slides, a separate page of instructions was passed out. This page explained the procedure for locating areas and listed the following nonsuggestive questions to be answered by the subject for each of his responses:

1. What characteristics of the blot suggested the particular thing you saw; that is, what characteristics do they have in common?
2. In what position did you see the thing you saw?

The responses to these questions were of aid in determining accurately the number of responses to each card.

The subjects for the experiment consisted of 238 students in General Psychology at Montana State University. The first tests were administered on April 6, 1948, to seven regularly scheduled discussion sections meeting from 8:00 A.M., to 4:00 P.M. Sections 1, 3, 5, and 7 were shown the regular Harrower-Erickson series, while sections 2, 4, and 6 were shown the achromatic series. At the second session, held on April 13, 1948, the sections responded to the alternate series.

[1]When all ten cards are photographed on a single plate with uniform lighting, the most noticeable feature of the resulting print is a marked gradation of brightness from Card I to Card X. The figures on the last three cards are so light that they can hardly be distinguished from their backgrounds.

TABLE I

SUBJECTS: DISTRIBUTION WITH RESPECT TO SEX AND COLLEGE CLASS; MEAN AGE AND MEAN OTIS S-A TEST SCORE (N = 238)

	Harrower-Erickson Series First	Achromatic Series First
Total N	129	109
Number of Males	87	66
Number of Females	42	43
Freshmen	82	72
Sophomores	27	24
Juniors	18	11
Seniors	2	2
Mean Age	21.0	20.5
Mean Otis S-A Score	59.3	59.9

In Table I are shown the distributions of subjects with respect to sex and college class; also shown are the mean age and mean Otis S-A Test score of the subjects. It will be noted that the subjects who were given the regular series first and those who were given the achromantic series first were similarly distributed with respect to these variables.

Since the aim of the experiment did not require full scoring, the only scoring procedure used was that of determining the number of responses to each blot. Each of the authors carried on this scoring procedure independently, using the standard criteria of scorable responses recommended by Klopfer [1, pp. 65-69] for individual administration. This independent scoring yielded 91.45 per cent agreement on the total of 4760 slides (20 slides for each of 238 subjects) for which responses were

counted. The percentage of agreement, when expressed in terms of 8-9-10 per cent, which has primary relevance for the present problem, was 99.49 per cent. In all cases of original disagreement, the decision was arrived at through discussion. Some of the disagreements, as might be expected, were attributable to recording errors during scoring.

TABLE II

COMPARISON OF HARROWER-ERICKSON SERIES AND ACHROMATIC SERIES WITH RESPECT TO MEAN PRODUCTIVITY OF SUBJECTS

	Harrower-Erickson Series	Achromatic Series
Mean R	16.9	17.2
Mean R (Cards 1, 4, 5, 6, 7)	8.33	8.33
Mean R (Cards, 2, 3, 8, 9, 10)	8.86	8.87
Mean 8-9-10%	31.84	31.40

RESULTS

The data of Table II provide comparisons of productivity of the subjects in response to the Harrower-Erickson series and to the achromatic series of Rorschach blots. It will be noted that mean productivity on the Harrower-Erickson series was highly comparable with mean productivity on the achromatic series. This was true for total productivity (Mean *R*), for productivity on blots ordinarily involving no color (1, 4, 5, 6, 7), for productivity on blots ordinarily involving some color (2, 3, 8, 9, 10), and for mean 8-9-10 per cent. It will further be noted that the mean 8-9-10 per cent on both the Harrower-Erickson series and the achromatic series closely approximated 30 per cent. All the mean differences were too small to be considered statistically significant. There seems, then, to be nothing inherent in the achromatic series which would tend, for the group as a whole, to depress or stimulate either general productivity or specific blot productivity below or above the level of productivity occurring when color is present on some of the blots.

TABLE III

COMPARISON OF HARROWER-ERICKSON SERIES AND ACHROMATIC SERIES WITH RESPECT TO DISTRIBUTIONS OF 8-9-10% (N = 238)

	Frequency	
8-9-10%	Harrower-Erickson Series	Achromatic Series
56-59	2	1
52-55	1	1
48-51	2	2
44-47	5	6
40-43	17	17
36-39	38	27
32-35	38	45
28-31	67	67
24-27	43	43
20-23	18	19
16-19	5	7
12-15	1	3
8-11	0	0
4-7	0	0
0-3	1	0
Standard Deviation	7.24	7.04

The crucial test of the hypothesis involves a comparison of the variability in 8-9-10 per cent among subjects on the Harrower-Erickson series with variability on the achromatic series. It has been pointed out above that if color is the factor responsible for variability in 8-9-10 per cent, then this variability should be reduced to a minimum when achromatic blots are involved in cards VIII, IX, and X as well as in the remaining cards; the proportion of responses to the last three cards should not vary appreciably from its expected value of 30 per cent. However, the data of Table III show that the distribution in 8-9-10 per cent on the achromatic series was highly similar to the distribution on the Harrower-Erickson series.

The magnitudes of the standard deviations were also essentially equivalent. When these distributions were tested for similarity by means of Chi-square, the Chi-square value was found to be 12.70. This value is sufficiently small that it would be expected to occur in between 10 per cent and 20 per cent of samples drawn from a population in which the two distributions were identical. The similarity which occurred in the variabilities for these two distributions, then, requires an interpretation in terms contrary to our hypothesis. Since variability in 8-9-10 per cent was approximately as great when completely achromatic blots were employed as when the Harrower-Erickson blots were employed, it requires that variability in 8-9-10 per cent be attributed to some other factor than responsiveness to color.

DISCUSSION

Since the evidence requires interpretation of variability in 8-9-10 per cent in terms other than the operation of individual differences in responsiveness to color, it seems reasonable to suggest alternative hypotheses. Three such hypotheses occur to the authors: (1) that variability may be attributed to low reliability of the 8-9-10 per cent; (2) that variability in 8-9-10 per cent may result from the fact that Blots VIII, IX, and X differ from the other blots in configuration; (3) that variability in 8-9-10 per cent may be ascribed to individual differences in responsiveness to the differential figure-ground contrast values of the different blots. The design of the present experiment was not of a nature to provide dependable evidence to confirm or deny the validity of any of these alternative hypotheses.

SUMMARY

The Harrower-Erickson series of Rorschach slides and a comparable series of achromatic Rorschach slides were presented to 238 college subjects, to determine whether variability in 8-9-10 per cent can be attributed to individual differences in responsiveness to color. Evidence was obtained which was contrary to the hypothesis, in so far as it applied to the group Rorschach administered under the conditions described. Whether similar results would be obtained when Rorschach blots are individually administered remains a problem for further investigation. Alternative hypotheses were suggested as possible explanations to account for variability in 8-9-10 per cent. The degree of validity possessed by these hypotheses also requires experimental investigation.

Received November 22, 1948.

REFERENCE

1. KLOPFER, B., AND KELLY, D. M. *The Rorschach technique.* Yonkers: World Book Co., 1942.

PROBLEMS OF COLLABORATION BETWEEN THE CLINIC CENTER AND THE PSYCHOLOGY DEPARTMENT IN PROVIDING CLINICAL EXPERIENCE

BY FLORENCE R. DAY
SMITH COLLEGE SCHOOL FOR SOCIAL WORK

CLINICAL psychology and psychiatric social work approach the problem of collaboration between the academic institution and the clinic center from opposite origins. The impetus for professional education in social work arose from social agencies and from social workers who felt inadequate to meet the day by day problems presented by their clients. The psychologist, deeply oriented in the academic, reaches out for an experience with human material to acquire aptitude in the application of what has been learned theoretically. In both approaches the problem of collaboration is the same, namely, to differentiate *clinical training* that may properly be of such educational stature as to deserve academic recognition from *clinical experience* that is merely learning by doing without educational controls.

Before discussing procedures in collaboration that have been developed in psychiatric social work and that may in some of their aspects be transferable to the psychologist's situation, certain underlying principles should be stated upon which these procedures have been predicated. The curriculum in professional education for social work envisages a systematic planned sequence of learning that combines academic study and supervised practice. The educational institution assumes full educational responsibility. This presumes that the university is as concerned with the subject matter, the teaching methods, and the progress of the student in the clinic setting as it is with the quality of the academic courses. It is one curriculum. It is one faculty, composed of classroom instructors and field instructors even though the latter are employed staff of the clinic. It is a continuous process of learning in which theoretical knowledge and practical methods are taught in the academic courses and theoretical knowledge and practical methods are taught in the clinic. This point of view is opposed to the conception that the teaching of theoretical knowledge should be confined to the classroom and the teaching of practical methods should be confined to the field. Instead, a course of instruction that alternates periods of study and application teaches and reteaches the theoretical and practical throughout each successive period as understanding deepens. In this way knowledge is translated into skill, and integration is achieved.

Furthermore it is assumed in professional education that the educational base must be a broad one covering a common core of subject matter and basic skills in order that the professionally trained person may be equipped to use his preparation in a wide variety of clinic settings rather than directing the learning toward a single type of service or program. Social work education arising as it did from agency need, has had a

struggle to break away from agency-oriented learning that resulted in too early and too much specialization. The major task of its educators has been to abstract from cumulative experience this educational core. This mistake you are in a position to avoid, but it is likely that you will find the university with its educational goal and the clinic center with its practical needs pulled in opposing directions.

What is implied when we say that the university assumes educational responsibility? First of all it implies clarity in the objectives as to what the student is to be trained for. That is, what he will need to know, what he must be able to do, and I would also add, what kind of a person he must become. The amount of emphasis given to the self-development of the student will be governed of course by the kind of activity he will be expected to assume. But whenever professional practice is anticipated, the attainment of desirable attitudes and emotional responses in the practitioner cannot be overlooked. The modification and maturing of the student's personality may be left to chance, or may be anticipated and directed as part of the educational objective.

To assure high educational standards in clinical training it is incumbent on the university to develop certain types of criteria that will serve partly as educational safeguards and partly as goals toward which university and clinic center together will strive. Criteria will need to be formulated for the selection of students because potential professional practitioners must possess the particular personal qualifications and aptitudes on whom this kind of learning will take. Criteria will also be needed for the selection of clinic centers, for the selection of field instructors or student supervisors within those centers, and for measuring the progress and proficiency of the student at specific periods in his clinical training. Sufficient and continuing attention to these matters requires constant revision and improvement, as such criteria are never static.

In saying that the university assumes full educational responsibility it is not my intent to imply that the clinic center has no educational responsibility. Full participation on the part of the field centers in the developing of criteria will assure realistic planning and the execution thereof.

The first prerequisite in selecting a clinic center is to be sure that the educational responsibilities that it is assuming are thoroughly understood and accepted. The clinic center must be committed to carrying on a teaching program as one of its major functions. This commitment should come from the top down and should not try to find its way via enthusiastic student through willing supervisor outward to encompass staff and upward to eventually reach the administrative heads. The governing body, whether it be commissioners, boards of trustees, or the so-called hospital or school authorities, should understand as fully as possible what will be expected, because the program will involve a drain on staff resources, on finances, on physical equipment, and it will have an effect on services rendered to its patients. The staff, from the executive to the newest clerical assistant, should give more than verbal approval, because each will play a role in the education of the student.

Clinical training is not confined to what takes place between the student and his individual cases, or to what takes place between himself and his supervisor. Obviously in a clinic setting the professional members of the clinic team (the psychiatrist, the psychiatric social worker and other personnel) must consciously supplement the teach-

ing role of the supervisor. But, more than this, clinic training does not isolate and emphasize one aspect of the curriculum for application, namely, diagnostic and therapeutic skills. The clinic center also offers an opportunity for the student to participate in the overall clinic program and in the community activities of the agency. Thus the clinic becomes a laboratory setting for applying theoretical and practical concepts in aspects of administration and community organization processes.

The second prerequisite for the selection of the clinic center has to do with the quality of practice performed throughout the agency. The student's learning is accelerated by having an opportunity to see practice performed at a level beyond his own abilities. Staff seminars, conferences on cases, individual consultation with staff members, are necessary aids to the tutorial system of teaching that we call supervision. Due to the dearth of qualified agencies accessible to graduate schools of social work, the device is sometimes employed of setting up a training unit under a qualified supervisor who will attempt to teach practice that is either above the level of the agency practice or outside the prescribed function of the agency. This procedure deprives the student of the concomitants provided through the total environment.

The prerequisite that I am stating as third should come first in order of importance. This is the selection of the student supervisor, or field instructor. Integration is primarily a matter of comparable quality on the part of class and clinic instruction. Although the clinic may designate staff members to carry the supervisory role, the university must participate in the selection and in the training of the supervisors. It is necessary that the supervisor be thoroughly competent in the kind of practice that it will be incumbent on him to teach; and the supervisor must be competent as a teacher in the use of appropriate teaching methods. If the supervisor is to carry out his share of the educational responsibility as defined above, he must also be thoroughly familiar with the total curriculum and able to teach theory as well as practice in the several areas of the curriculum. If he has not had the advantage of the full training that is planned for the student, it is incumbent upon the university and upon him to fill in the gaps. Schools of social work, together with field agencies, undertake quite an elaborate program of extension courses, seminars, informal group discussions, and individual consultation that aim to keep the supervisor currently informed regarding class content and that devote time to discussion of supervisory problems and methods.

The last prerequisite that I will mention refers to practical arrangements. It is necessary that the work schedule of the supervisor who carries the double role of staff member of the clinic and faculty member of the university be planned to permit sufficient time for uninterrupted conferences with the student. "Sufficient" needs to be clearly defined in terms of length and frequency. In order to insure against interruption it has been found inadvisable for the supervision to be undertaken by a person who has a major administrative function. The clinic exists for service to the patient and to the community, and interruptions from these sources deserve precedence from the administrator. The educational program exists for the student and the student should have priority in the meeting of his educational needs.

At the risk of stressing the obvious I will stress the importance of other facilities, such as the availability of

library material. Reading undertaken in relation to practical case situations *clicks* so that even known material takes on new meaning and should be accessible. Privacy for interviews, whether with the supervisor or the patient, directly affects the quality of those interviews. The initiative taken by the student for planning conferences with client or supervisor, for preparing material for staff discussions, for study and evaluating his own performance, will be furthered by a physical setting that gives him a desk, a private room if possible, and prompt and efficient stenographic service.

The enumeration of these prerequisites for the selection of the clinic center indicates that the university department must not only establish criteria but must have procedures for determining whether the clinic center actually possesses these attributes or will take the necessary steps to develop them. This kind of evaluation should take place before an affiliation with a clinic center is confirmed. We see this as involving a visit to the center by the director of the program, who meets with the administrator, the staff, and desirably, the governing board. It also involves a visit by the faculty member who may and may not be the director who is responsible for evaluating the quality of the agency practice and who will participate in the selection of the supervisor. Opportunity to attend staff discussions, to read clinic records, to confer with professional members of the clinic team, afford first-hand evidence on which such evaluation may be made.

When the standards that will be expected of the clinic are formulated and presented in a direct fashion the clinic's respect for the school is increased, and thereby its own feeling of prestige. We find that agencies like to have students. They like the prestige of being known as a teaching agency and of having university affiliation.

When the selection of centers need not be confined to the campus setting or the immediate community, as is possible when the practicum periods alternate with periods of academic study, it enables the university to utilize agencies over a wide geographic range that possess uniform standards in respect to these criteria. It also affords a wider variety in the types of clinics used, so that the student may be assigned to a placement particularly adapted to supplement and complement his previous training and experience—a tailor-made placement, if you please.

Once the clinic center is affiliated and the students are assigned, the university must put into operation methods for the continued supervision of the clinic training. At the Smith School this consists of a complex plan for regular visits by members of the School faculty. The Director visits each agency once each year. It is important that the director of the program be thoroughly familiar with each center in its entirety, and more specifically with details of the student program. Each student is assigned a faculty advisor who is responsible for the student's progress in his practical work and a thesis advisor who is responsible for the student's research project. The faculty advisor and the thesis advisor visit each student at least three times during the eight-and-one-half month period in the clinic. The faculty advisor knows the student in his academic work and is identified with the procedures and policies of the university. The faculty advisor talks separately with the field instructor and the student and consults with others in the agency who share responsibility for the student program. These visits are not conceived as being limited to a checking procedure or even to the giving of constructive

criticism. The faculty advisor discusses freely with the supervisor and with the student what is happening in the educational process. Because the faculty advisor has a wider perspective through knowing a number of students and clinic settings, he can be helpful to both the student and supervisor in removing barriers to learning, in working out difficulties that may arise, in making suggestions regarding educational content, and in integrating the total educational experience. It is important that the faculty advisor should not come between the student and the supervisor, but should handle any difficulties that arise in a way to strengthen this relationship. It is our experience that these visits of the faculty advisor mobilize both student and supervisor so that spurts in growth can be definitely distinguished following the visits. These visits are the surest method for facilitating integration between learning in the classroom and learning in the field.

At Smith School where our students are placed at a distance, certain other methods have been devised to facilitate integration and assure a high quality of learning during the practicum period. The clinic supervisor sends monthly progress reports to the School and makes a formal evaluation in the middle, and at the end of the eight-and-a-half month field work period. The School finds these monthly reports invaluable in keeping in touch with students' progress. Supervisors in turn find that having to write a statement about each student enables them to catch many trends or gaps in the students' experience that they might otherwise have become aware of much more slowly.

Students send two types of reports to the School every month. One is primarily a statistical statement indicating the number of cases carried, supervisory conferences held, and staff conferences and meetings attended. They also write a six-line statement of their social diagnosis and plan of treatment for each of their new cases. The supervisor countersigns this report and thus sees at a glance how well the student's experience is balanced.

The other report which students send directly to the faculty advisor is based primarily on reading. Students are required to read at least six hours a week throughout the winter. Their reports are not book reviews. They are asked, rather, to discuss their reading as it relates to their experience. These reports in the beginning are usually rather externalized, but as time goes on they reflect the student's increased awareness of his own part in the learning situation.

The student who has gone through a long period of academic preparation comes fearful and apprehensive of the testing ground because he realizes the seriousness of undertaking to apply his knowledge to human lives. The supervisor must be prepared to handle this initial psychological problem. Some students meet this situation by an almost complete lapse of memory regarding the academic content. Others go to the opposite extreme of giving off a rash of theory which acts as a barrier between themselves and their supervisor and also between themselves and the patient. The whole clinic setting may be unfamiliar, threatening, and bewildering. If it is an extensive agency offering multiple services it is even more difficult for the trainee to find his place. The supervisor becomes his first source of security. When the student has established a comfortable relationship with his supervisor he can begin to take in the horizon around him. We find the problem is solved better in this way than by plunging the student into a planned orientation program at once. He *does* need to feel welcome, which can

be handled by introduction to the people with whom he will first come in contact; by having a place ready for him in which he may work and feel is his own; and by acquainting him with the facilities that he will need to use in his own activity during the first few days. The supervisor should know something about the student before he comes, should certainly be furnished with the so-called vital statistics, his educational and experience background and his progress as revealed in the course of study. This is essential in order that specific work assignments be ready for him to undertake. There is nothing like a work assignment, providing it is within the limits of successful accomplishment, to allay the initial anxiety.

During the period of practical training the trainee is both worker and student. This is a difficult concept to keep in balance. Through real participation in the real work of the clinic the student attains professional stature, but at the same time his status as a student must be jealously guarded. To distinguish again between clinical training and mere clinical experience, clinical training puts first the needs of the student, second the needs of the patient, third the needs of the clinic. Cases are deliberately selected in which the needs of the patient are somewhat more complex than the ability of the student to meet them, but this is a necessary incentive to learning. The patient is protected by the supervisor, who assumes responsibility for the student's performance. This is in contrast to mere experience in which the need of the agency comes first, the need of the client second, and the need of the worker last. In this instance patients are selected to match the ability of the worker in order that the least harm will be done to the patient, and the maximum service in quantity be rendered in the clinic. The agency in conducting the training program must needs adopt the long view of sacrificing more immediate returns in amount of work at minimum quality to the acquisition of depth and breadth in the performance of future personnel. The student's own need to feel useful and to be busy may be readily exploited. Selection of cases according to the increasing competence of the student will make it possible for the student to function as a productive staff member with accruing satisfaction both to the student and the clinic.

The student-supervisor relationship is the element that above all others supplies the dynamic for the kind of learning that encompasses the emotional component as well as the intellectual in professional growth. For this to happen the supervisory relationship must be developed in such a way that the student is free to discuss his likes and his dislikes, his belief and his doubts, his identifications and his biases as he takes on new ideological concepts and as he experiences the use of himself in new situations. Dealing with human material brings to the fore the emotions as well as the mind and in the student as well as in the patient. Thus the supervisor gains an increased understanding of the blind spots, the prejudices, the emotional blockings that get in the way of successful work. Personal difficulties in the student may also come to the fore and the supervisor is responsible to be helpful with these only insofar as they affect the learning situation.

If self-development is to be directed rather than left to chance, continuity in the student-supervisor relation is essential. We find that students enter the practice period with high enthusiasm for their cases and with a beginning self-confidence. As they become sensitive to the intricacies of working with human beings the ups and downs set in

—the downs taking the form of discouragement, doubts and apprehension regarding their ability. Toward the end of the period, if it is of sufficient length, a kind of leveling off takes place, when a proper synthesis of competence and confidence is reached.

The rate of learning varies tremendously with individual students and the criteria for evaluation must be kept flexible. For these reasons we find that continuity with one supervisor in one setting brings better results than frequent rotation or change.

At the end of a second field work period a student wrote in her last month's report to the Faculty Advisor as follows:

> The year before I came to Smith I took a course in casework. The assignment was to read Gordon Hamilton's *Theory and Practice of Social Casework.* As it was one of the first assignments in the field I was quite conscientious and chapter by chapter I outlined the book. The neatest printing job I've ever done. After memorizing the notes, I carefully put them away and the other day came across them for the first time in over two years. It was absolutely amazing to me to read them again. It was like reading them for the first time, and yet I remember learning them by heart . . . . or maybe it was by rote, not by heart.

In summary, it is our belief that without the period of clinical training professional growth and competence in the practitioner will not take place. Only through one's own activity can one acquire knowledge in the vital way that is necessary if it is to be applied in new situations. Only by actually taking responsibility for patients will professional self-discipline and self-reliance be developed.

Received December 23, 1948.

AN EVALUATION OF WECHSLER-BELLEVUE PATTERNS IN SCHIZOPHRENIA

By SOL L. GARFIELD
UNIVERSITY OF CONNECTICUT

PSYCHOMETRIC patterns and analyses of scatter have been used for some time in psychology as diagnostic devices [2], and were given renewed impetus with the publication of the Wechsler-Bellevue Adult Intelligence Scale [11]. Since the W-B scale consists of eleven subtests and employs a system of weighted scores which allows intersubtest comparisons, it lends itself easily to pattern analyses. A variety of different techniques for evaluating subtest patterns has been offered by different investigators. The greatest application appears to be with patients diagnosed as schizophrenic. Some investigators have reported diagnostic patterns for such patients which have been based on the mean subtest scores for an entire group of subjects. In other words, the pattern is based on the mean performance of a group rather than on the patterns of individual subjects analyzed separately. The usual procedure in such an approach is to compare the rank order of subtest means from a group of schizophrenics with the order of subtests derived from normal or control groups [8]. Such an approach overcomes the difficulty of comparing groups whose absolute scores differ so much that comparison of actual scores would be somewhat meaningless. On the other hand, some investigators merely compare the mean subtest scores of schizophrenics with those of normals and where statistically significant differences between subtest scores are obtained, a diagnostic difference is inferred [7]. This method, of course, overlooks the fact that the difference in IQ or mental level between the schizophrenic and normal groups affects the mean subtest scores.

In addition to patterns derived from the mean scores of groups, patterns have also been reported in terms of an individual's subtest variation from some fixed point. Some measures are based on the deviations of subtests from a subject's own mean [3, 5], while others rely on deviations from the verbal or performance mean [11] or from the vocabulary level of the subject [10]. Although the overall methods of analysis and treatment of data vary, the latter methods are based on deviations or patterns of individuals rather than on mean measures of a group. For this reason, they more closely approximate the actual clinical use of such measures since the individual is the point of reference.

Since, among other things, different techniques have been used for deriving psychometric patterns in schizophrenia, perhaps some lack of agreement between the various investigations is to be expected. In a previous paper [3], the author reported an evaluation of certain existing diagnostic patterns for schizophrenics. In that study, the actual frequency of subtest deviations from each subject's mean was used as the basis of evaluation and a nonpsychotic hospitalized control group was used for purpos-

es of comparison. The major findings were: (1) The failure of most diagnostic patterns to adequately differentiate schizophrenics from control subjects; (2) the large amount of overlap in subtest deviations between schizophrenics and controls; and (3) the variability in patterns within the schizophrenic group. Certain hypotheses were offered to explain these findings as well as the conflicting findings reported in other studies. Among these were such factors as the lack of comparability of schizophrenic subjects in the various studies, differences in diagnostic criteria, variability of subjects in terms of degree of disturbance, and problems of motivation.

The present study[1] is in part an outgrowth of the previous one. However, whereas the former study was primarily concerned with the diagnostic efficiency of test patterns in the clinical situation, the present investigation is an attempt to explore more general problems and assumptions underlying the psychometric pattern approach to diagnosis in schizophrenia. Special attention is paid to certain sampling problems which heretofore have received little emphasis. The general plan of the paper will be to analyze first the patterns of mean subtest scores for an entire group of schizophrenic subjects and to compare them with the findings of other studies, to evaluate the influence of such factors as education and IQ on psychometric patterns, and finally, to discuss the individual variability of patterns.

[1]The writer acknowledges the aid of Mr. William Barclay in the tabulation and analysis of the data in this study.

SUBJECTS

The subjects used in the present investigation were 109 male patients in a veterans hospital. All were diagnosed at staff conferences as schizophrenic. The sub-group classifications were as follows: unclassified or undifferentiated, 36; paranoid, 23; hebephrenic, 20; simple, 9; catatonic, 9; mixed and others, 12. The average age of the subjects was 31.21 (SD, 10.81). The average grade reached in school was 10.43 (SD, 3.06) and the mean full scale Wechsler-Bellevue IQ was 98.63 (S.D., 17.53). The test records were obtained during consecutive test administrations in the hospital and the only selective require-

TABLE I
MEANS, SIGMAS AND RANK ORDER OF W-B SUBTESTS IN FOUR STUDIES OF SCHIZOPHRENICS

	Garfield			Gilliland et al.			Magaret			Olch		
N	109			87			80			32		
Mean Age	31.21			33.7			35.7			32.28		
Mean IQ	98.63			87			87*			89.57		
Subtests	Mean	SD	Rank	Mean	SD	Rank	Mean	SD	Rank	Mean	SD	Rank
Info.	10.06	2.84	1	7.31	3.02	5	8.1	3.4	1	8.88	2.07	3
Compr.	9.29	4.07	5	7.78	3.44	3	6.0	3.5	6	7.03	2.42	7
Arith.	8.69	4.56	7	6.98	7.14	7	5.4	3.9	8.5	7.25	3.37	5
Digits	8.03	3.41	8	7.14	3.24	6	6.6	2.6	5	8.22	3.24	4
Simil.	9.12	3.98	6	7.79	3.32	2	6.8	4.0	4	6.38	2.13	9
P.C.	9.37	3.26	4	7.70	3.21	4	5.8	3.7	7	6.91	2.80	8
P.A.	7.92	3.36	9	6.76	3.12	9	5.4	3.2	8.5	7.09	3.27	6
O.A.	10.00	2.83	2	6.96	3.44	8	7.4	3.6	2	9.47	3.32	1
Bl.D.	9.67	2.96	3	8.44	4.07	1	7.2	3.5	3	8.91	2.74	2
Dig. Sym.	7.72	2.68	10	6.14	2.91	10	5.2	2.9	10	6.00	1.91	10

*Estimated Mean IQ in Terms of Average Age of Subjects. Mean IQ Not Given in Original Publication.

ment used for inclusion in the study was that the subject complete all the subtests of the Wechsler-Bellevue scale.

RESULTS

Since several other studies have reported scatter patterns in terms of mean subtest scores or the rank order of subtests within a given nosological group, these procedures will be used in order to facilitate comparisons. It is emphasized, however, that this procedure is feasible only for group comparisons, and cannot be taken to indicate diagnostic patterns for the individual case.

The subject's mean subtest scores along with the sigma and rank order of the subtests are given in Table I. Also included in this table are comparable data from three other studies, [4, 5, 7]. It is apparent from the mean IQ as well as the subtest means that the schizophrenic groups differ in terms of the level of mental functioning. The subtest means for the present group of subjects, for example, are considerably higher in most cases than the means of the group tested by Magaret [5]. It is for this reason that a ranking method has been used in other studies [7, 8] to compare the patterns of different groups of subjects. A comparison of the rank order of the subtests shows some similarities in pattern in the different studies as well as some important differences. Whereas the Digit Symbol subtest is consistently last in rank in all four studies, and the Block Design subtest varies only within the first three ranks, the order of other subtests is not as consistent. Lack of consistency is particularly apparent in the ranks accorded the Object Assembly and Similarities subtests which range from 1 to 8 and 2 to 9, respectively. Other discrepancies are not as marked. However, if comparisons are extended to include other studies [8, 12], additional variations in the ranking and patterning of subtests are found. The size of the sigmas of the subtests is of interest in indicating the variability of performance within each of the groups studied, as well as differences in subtest variability between the groups in the four studies.

The differences in diagnostic patterns secured in the various studies have been emphasized in several recent reports [3, 7, 9] and need only brief mention here. The pattern variations are perhaps better seen graphically as portrayed in Figure 1. It is interesting that such variations exist since the patterns are based on mean scores of different samples of schizophrenics. One would normally expect patterns derived from groups to be relatively stable. One probable explanation for such discrepancies is that the various schizophrenic samples are not comparable in many respects. An assumption usually made in such studies is that one group of patients diagnosed as schizophrenic is comparable to another group of patients bearing the same diagnosis. Obviously, this is not so.

In most instances, subjects selected for study have been selected primarily on the basis of diagnosis. The reliability of such diagnosis is of course a central problem in psychopathological research. In addition, there are other sampling variables which have received rather scant attention. One such variable, however, the age of subjects, has received consideration in several studies[1, 7, 9, 12]. The latter have shown some differences in patterns between schizophrenic groups of different ages. Some diagnostic patterns have been found to apply to younger schizophrenics, but not to older groups. This finding alone limits the value of psychometric patterns derived from groups where age differences are overlooked. Furthermore, such factors as sex, education,

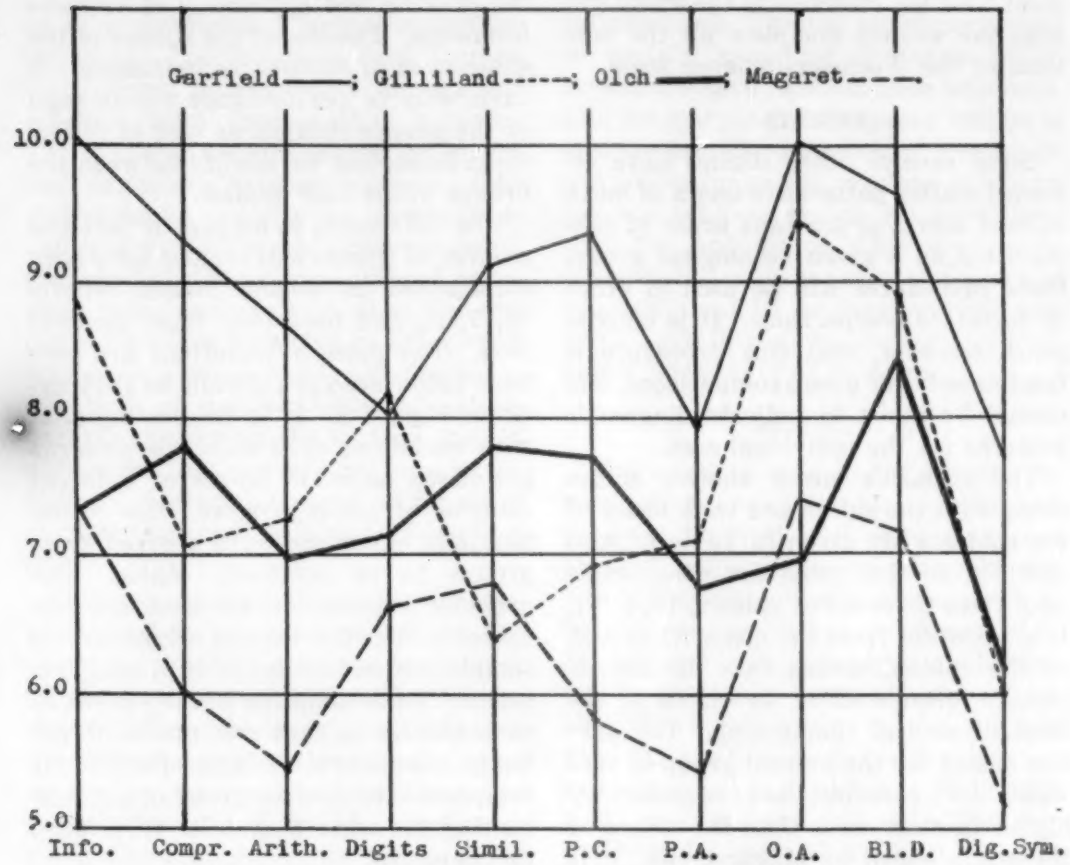

FIG. 1. Patterns of mean subtest scores for Schizophrenics in four studies.

level of IQ and severity of disturbance have been slighted in investigations of diagnostic patterns for different clinical groups. Although the importance of these factors are recognized [2, 9], few investigators have made allowances for them in their research projects. Rapaport [10] defends this lack in his own work on the assumption that deviations from an individual's mean based on weighted scores rule out such problems. Similar assumptions apparently are made by others for little attention has been given to such matters. Although education level, IQ and socio-economic status are considered in research on normative groups, it would appear that once an individual is given a psychiatric diagnosis, such considerations are of little or no consequence. Nevertheless, since such factors are known to have some influence on verbal test performance, it is worthwhile to analyze scatter patterns of schizophrenics in terms of such variables.

SUBGROUP ANALYSIS

To evaluate the influence on psychometric patterns of some of the factors mentioned above, the data secured from the entire 109 subjects were reanalyzed in terms of age, educational level and IQ. The subgroupings used were simple and conventional classifications. For example, the group was subdivided on the basis of three educational categories:

elementary school (up to grade 8), high school (grades 9-12) and college, (grade 13 and above). Similar classifications were used for IQ (below 90, 90-109, 110 and above) and age (17 to 29, 30 to 39, 40 and above). In addition, separate means were computed for the various schizophrenic subclassifications; i.e., paranoid, hebephrenic, etc. Representative data for groups separated on the basis of education are given in Table II. Since there were such wide differences in the actual weighted scores obtained, the data for all groups are not given. Instead, the various subtest means were ranked for each group and these findings are listed in Table III for all of the various subgroups analyzed.

TABLE II

MEAN IQ'S AND SUBTEST SCORES FOR SCHIZOPHRENICS GROUPED IN TERMS OF EDUCATION

Subtests	Grades 8 & Below N—34	Grades 9-12 N—53	Grades 13 & Above N—20
Ver. IQ	86.59	99.17	119.60
Perf. IQ	87.23	99.81	110.40
Tot. IQ	86.11	99.67	116.85
Info.	7.80	10.40	13.15
Compr.	7.27	9.58	12.25
Arith.	5.71	8.73	13.35
Digits	6.32	8.02	11.15
Simil.	6.76	9.11	13.20
P.C.	7.00	9.96	11.90
P.A.	5.71	8.51	10.20
O.A.	8.91	10.23	11.15
Bl.D.	7.50	10.42	11.30
Dig.Sym.	5.79	8.30	9.80

Although some of the groups are small, it is apparent that there are some important differences between the various groups. Probably the most striking difference is found on the Arithmetic subtest when educational level is the basis of comparison. As might be expected, the lowest educational group performs at its lowest level on this subtest whereas the group with some college education obtains its highest score on this subtest. Other striking differences in rank are noted on the Object Assembly subtest when education is considered and on the Similarities subtest when level of IQ is taken into consideration. Comprehension is another subtest where meaningful differences are also observed between the various groups, particularly between paranoids and hebephrenics. The most consistent subtest is Picture Arrangement on which all groups perform relatively poorly. This cannot be interpreted as a diagnostic index for schizophrenia, however, for other pathological groups may also

TABLE III

RANK ORDER OF SUBTEST MEANS FOR VARIOUS GROUPINGS OF SCHIZOPHRENIC SUBJECTS

Sub-Tests	Entire Group N–109	Age Groups 17-29 N–62	30-39 N–24	40 & Above N–23	Educational Groups Grade 8 & Below N–34	Grades 9-12 N–53	Grades 13 & Above N–20	IQ Groups 89 & Below N–33	90-109 N–45	110 & Above N–31	Diagnostic Subgroups Uncl. N–36	Par. N–23	Heb. N–20	Sim. N–9	Cat. N–9	Others N–12
Info.	1	3	1	1	2	2	3	2	1	3	1	2	2	5	4	2
Compr.	5	5	4	4	4	5	4	7	2	4	4	1	9	3	6	8
Arith.	7	7	7	5	9.5	7	1	10	7	2	7	5	5	9	6	6
Digits	8	9	10	7	7	10	7.5	6	8	8	10	9	4	10	6	7
Simil.	6	6	3	6	6	6	2	9	5	1	6	3	6	4	8	3
P.C.	4	4	5	8	5	4	5	4	6	5	3	4	7.5	6	3	1
P.A.	9	8	9	9	9.5	8	9	8	9	9	8	8	10	8	10	9
O.A.	2	1	2	2	1	3	7.5	1	3	7	2	7	1	2	1	4
Bl.D.	3	2	6	3	3	1	6	3	4	6	5	6	3	1	2	5
Dig. S.	10	10	8	10	8	9	10	5	10	10	9	10	7.5	7	9	10

perform poorly on it [3]. One of the most important facts to be noted in these data is the disparity between the rank of a subtest for a specific subgroup and the rank derived for the subtest from the entire group of subjects. This is clearly seen in the case of Arithmetic, Similarities and Object Assembly. The relative position of some subtests thus appears to be closely related to the education or IQ of the subject. This, of course, limits the effectiveness of group diagnostic patterns. For example, it has been reported by some investigators [8, 11] that the verbal score exceeds the performance score in schizophrenics. Other studies, however, fail to confirm this [3, 7, 9]. Perhaps some light can be thrown on this matter by the subgroup analyses. In Table II it is noted that the mean Verbal IQ for the lowest educational group is slightly below the mean Performance IQ. The opposite is true for the highest educational group. Furthermore, when the individual records of each patient in the different IQ groups is examined for this pattern, the trend is even more distinct. Whereas 21 out of 31 subjects in the highest IQ group had higher verbal scores than performance scores, only 10 out of 33 had such a pattern in the lowest IQ group. It is apparent, therefore, that superiority of the verbal score over the performance score is related to the educational and IQ level of the subject. Patterns derived from a large heterogeneous group of subjects are obviously not applicable to individual cases or specialized subgroups.

Such findings cast serious doubt on the efficacy of comparing the subtest means of two groups which differ markedly in IQ. The mean mental level of the groups compared will affect not only the mean subtest scores but the patterning of such scores as well. Nevertheless, a recent paper [7] compares the mean scores of schizophrenics with those of normals in order to find significant differences in subtest performance. The fact that there was a difference of approximately 10 points between the mean IQ's of the two groups studied would appear to influence the significance of the differences found. In fact, if one compares the means of Gilliland's schizophrenic group with those for the group in the present study, statistically significant differences at the 1 per cent level of confidence are found for five of the subtests. If the present group is compared with Magaret's group, statistically significant differences ($P<.01$) between the means are secured for all 10 subtests. While the groups are not really comparable in age and other factors, some of the difference undoubtedly is due to the discrepancy between the groups in IQ. The findings reported in this section also do not lend support to Rapaport's contention that age, IQ and education are not important in diagnostic studies of scatter. Even when the basis of comparison is the individual's own pattern of subtest deviation from his mean score, the influence of such variables is evident. A majority of the patients in the lowest educational category have negative deviations on the Arithmetic subtest while 80 per cent of the patients in the highest educational category exhibit positive deviations on this subtest. Exactly the opposite relationship is noted on the Object Assembly subtest. On the latter, one-half of the highest educational group secures negative deviations as contrasted with one-seventh of the lowest educational group. In addition, the lowest educational group secures over twice as many positive deviations on this subtest as the superior education group. There is, therefore, a definite relationship between the scores on certain subtests and the educational level of the subjects.

INDIVIDUAL VARIABILITY

Thus far, the discussion has centered around group patterns of schizophrenic subjects as well as patterns derived for specially selected subgroups of such subjects. The latter analyses were undertaken to evaluate the importance of such variables as age, education and mental level on the psychometric patterns of schizophrenics. Although the means of subtests were used as the primary basis of comparison, the findings indicated some important differences between subgroups of schizophrenics separated in terms of such variables. The data point up the limitations of current approaches which overlook such factors in their search for group diagnostic patterns. However, while such factors need to be considered in the design and evaluation of research, the individual still remains the important unit in clinical work. For functional effectiveness, diagnostic indices and patterns must be evaluated in terms of the individual.

In order to evaluate individual patterns on the W-B scale, each subject's subtest deviation from his own mean score was used. This technique is the same as that described in the previous report [3], and is a measure of the subject's variability about his own mean. The frequencies of individual subtest deviations for the entire group of subjects are summarized in Table IV. While some of the subtest patterns relating to positive and negative deviations are similar in trend to the mean scores reported in Table I, the variability of the deviations is impressive. There is practically no subtest on which all schizophrenics perform in a uniformly superior or inferior manner. Even when separate analyses are made in terms of educational level, individual variability is still present. Furthermore, although many significant negative deviations are observed for such a test as Digit Symbol, the trend cannot be interpreted as diagnostic of schizophrenia unless this pattern is found only for this nosological group. Previous findings along this line indicated similar trends for other pathological groups [3]. A diagnostic measure, therefore, must be evaluated in terms of its ability to differentiate between several clinical categories. For this reason, normal groups are not the most desirable for comparative purposes. When pathological groups are substituted for normal groups as con-

TABLE IV

PERCENTAGES OF SUBTEST DEVIATIONS ON THE W-B SCALE

N = 109

	Deviations							
	−3.50 and below	−2.50 to −3.49	−1.50 to −2.49	−.50 to −1.49	−.49 to +.49	+.50 to +1.49	+1.50 to +2.49	+2.50 to +3.49	3.50 and above
Info.	0.92	2.76	3.67	12.83	16.50	19.24	19.24	12.83	11.92
Compr.	6.42	2.76	10.09	17.46	12.83	15.58	17.46	13.76	7.34
Arith.	15.58	6.42	10.09	11.01	16.50	10.09	16.50	5.51	8.25
Digits	13.76	11.92	13.76	19.24	12.83	10.09	11.01	3.67	3.67
Simil.	3.67	7.34	11.01	13.76	16.50	22.95	15.58	2.76	6.42
Voc.*	1.89	4.72	9.45	24.52	17.91	21.68	11.32	2.83	4.72
P.C.	4.59	5.51	5.51	15.58	20.18	17.46	14.67	7.34	0.92
P.A.	13.76	11.92	15.58	23.93	7.34	11.92	8.25	6.42	0.92
O.A.	5.51	4.59	8.25	6.42	15.58	17.46	14.67	5.51	22.04
Bl.D.	0.00	5.51	8.25	12.83	15.58	22.04	21.10	8.25	6.42
Dig. Sym.	11.92	14.67	22.04	19.24	18.33	4.59	3.67	1.83	3.67

*Vocabulary scores were not secured for three cases and therefore on this subtest, N—106.

trols in diagnostic studies of schizophrenia, differences between groups usually diminish [6, 8]. The data in Table IV are presented to show the variability of scatter within one schizophrenic population. They do not provide a basis for diagnostic patterns since comparable data for other clinical groups are lacking.

SUMMARY AND CONCLUSIONS

Data secured from the Wechsler-Bellevue records of 109 schizophrenic patients have been analyzed and evaluated in terms of psychometric patterns. To explore the problem from several different vantage points, various methods of analysis were used. The mean subtest scores for the entire group were ranked and compared with other similar studies. While there is consistency in the ranking of such tests as Digit Symbol and Block Design in most studies, discrepancies in rank are also noted for such subtests as Object Assembly and Similarities. These differences in subtest patterns between the various studies may be due to the fact that the different groups of schizophrenic subjects were not comparable in many respects. With few exceptions, subjects are selected for study primarily on the basis of diagnosis. The limitations of such criteria plus neglected sampling problems undoubtedly contribute to the discrepancies found.

In order to check the importance of some of these sampling variables, the scores of the subjects in the present study were analyzed in terms of age, IQ, education and schizophrenic subclassification; i.e., paranoid, simple, etc. Important differences in subtest patterns were found when the entire group of schizophrenics were subdivided in terms of these groupings. With the exception of age, most of these factors have been neglected in previous work on diagnostic patterns. Since performance on the Arithmetic, Similarities and Object Assembly subtests particularly are related to level of education and IQ, patterns derived from a total schizophrenic group would appear to have limited value. Such factors must be considered in the design and evaluation of research.

The last part of the paper was concerned with the subtest deviations of each individual subject from his own mean. While the previously mentioned group and subgroup analyses have important implications for diagnostic research, the significance of any diagnostic measure must be evaluated in terms of its clinical value for the individual case. Such study of individual subtest deviations, similar to previously reported analyses, reveals considerable variation in the patterns of schizophrenic patients. There does not appear to be any clear cut psychometric pattern which can be labeled "schizophrenic." Where some trends seem to exist, comparable data for other psychopathological groups must be secured before anyone can consider the trends to be of diagnostic significance. In terms of present findings, therefore, one must be critical of attempts to devise and apply psychometric patterns based solely on broad nosological categories.

Received December 2, 1948.

REFERENCES

1. BALINSKY, B. An analysis of the mental factors of various age groups from 9-60. *Genet. Psychol. Monogr.*, 1941, 23, 191-234.
2. CARTER, J. W., JR. AND BOWLES, J. W., JR. A manual on qualitative aspects of psychological examining. *J. clin. Psychol.*, 1948, 4, 109-150.
3. GARFIELD, S. L. A preliminary appraisal of Wechsler-Bellevue scatter patterns in schizophrenia. *J. consult. Psychol.*, 1948, 12, 32-36.
4. GILLILAND, A. R., WITTMAN, P., AND GOLDMAN, M. Patterns and scatter of mental abilities in various psychoses. *J.*

gen. Psychol., 1943, **29**, 251-260.

5. MAGARET, A. Parallels in the behavior of schizophrenics, paretics, and presenile non-psychotics. *J. abnorm. soc. Psychol.*, 1942, 37, 511-528.
6. MAGARET, A. AND WRIGHT, C. Limitations in the use of intelligence test performance to detect mental disturbance. *J. appl. Psychol.*, 1943, 27, 387-398.
7. OLCH, D. R. Psychometric patterns of schizophrenics on the Wechsler-Bellevue Intelligence Test. *J. consult. Psychol.*, 1948, 12, 127-136.
8. RABIN, A. I. Test-score patterns in schizophrenic and non-psychotic states. *J. Psychol.*, 1941, 12, 91-100.
9. RABIN, A. I. The use of the Wechsler-Bellevue Scales with normal and abnormal persons. *Psychol. Bull.*, 1945, 42, 410-422.
10. RAPAPORT, D. *Diagnostic psychological testing.* Chicago: Year Book Publishers, 1945. Vol. I.
11. WECHSLER, D. *The measurement of adult intelligence.* (3rd. Ed.) Baltimore: Williams & Wilkins, 1944.
12. WEIDER, A. Effect of age on the Bellevue Intelligence Scale in schizophhrenic patients. *Psychiat. Quart.*, 1943, 17, 337-345.

RELIABILITY OF THE SHORTENED MINNESOTA MULTIPHASIC PERSONALITY INVENTORY

BY JULES D. HOLZBERG AND SALVATORE ALESSI
PSYCHOLOGICAL LABORATORIES
CONNECTICUT STATE HOSPITAL

CONSIDERABLE research with the Minnesota Multiphasic Personality Inventory (MMPI) has demonstrated that this psychological technique has validity in differentiating the strengths of psychiatrically significant personality components as well as in detecting invalidity due to confusion, poor comprehension or uncooperativeness [1, 2, 6, 7, 8, 9, 11, 12, 13, 14, 18, 20, 21]. Several studies have evaluated the significance of the pattern of the profile derived from the inventory in addition to the examination of peaks on single scales [6, 17, 21]. Although a favorable attitude toward the MMPI is not shared by all investigators, and although certain of its scales appear to have greater validity than others [2, 3, 4, 6, 19], it still remains one of the most useful nonprojective personality tests that can be used in the clinical setting. The inventory samples a wide area of significant behavior, is simple to administer and score, and is capable of providing supplementary psychiatric information which may have remained untapped in the course of ordinary psychiatric study of a patient. The test consists of 550 cards on each of which is printed a statement to which the subject is asked to respond in one of three ways: True, False, or Cannot say. At the time of this study, the test consisted of twelve scales[1] of which three were validating scales and nine were of diagnostic significance (hypochondriasis, depression, hysteria, psychopathic deviate, masculinity-femininity, paranoia, psychasthenia, schizophrenia, and hypomania.)

Of the 550 items, only 350 are actually used in the scoring. The 200 additional items constitute a reservoir from which the authors of the MMPI plan to construct additional scales for personality study. In current clinical practice, where there are large numbers of patients to be studied and limited staff to perform such studies, it is necessary to reduce testing time to a minimum. The use of these additional items, without minimizing their importance for future refinement of this inventory, nevertheless constitutes a drain on much needed testing time.

Ferguson [5] removed the cards not used in scoring because he found no marked difference in results by so reducing the inventory, but he offers no reliability data. Gough, on the other hand, in discussing the question of whether it is valid to reduce the tests to the actual items used in scoring states:

> It is questionable whether or not these 199[2] cards should be deleted in order to save time in administering and scoring the test, as is sometimes done. All of the questions contribute towards the context, or the "question-environment," about which the test was standardized. The influence of these 199 questions on a subject's responses to the 351 scored items is an unknown, but important factor, and until we know something of the effects of this inter-

[1]This present study does not include the new K scale [18] which was published after the completion of this study.

[2]Gough's count of 199 unused items differs from our own count of 200 items.

action we have no accurate idea of the validity of the norms when the test is shortened in this way. [6, p. 36]

The present study was undertaken to investigate the reliability of reducing the MMPI to the items actually used in scoring (the short form) in comparison to the whole inventory (the long form), using both clinical and statistical approaches.

PROCEDURE

The short form of the MMPI was prepared by eliminating the 200 items not used in the scoring. The procedure employed was to test psychiatric patients with both the long and short forms shortly after their admission to a psychiatric hospital. The patients were not selected according to any preconceived plan. Thirty more or less consecutively admitted patients were chosen as subjects, the only criteria being sufficient reading ability for understanding the statements, and the ability of the patient to give a representative performance with good attention, cooperation, and effort. One half of the group was tested with the long form first and within one or two days was retested with the short form. This procedure was reversed for the other half of the group. The psychiatric status was unchanged for each of these subjects during this one to two day period. The records were scored according to instructions in the MMPI Manual [10] before the K scale had been added. Correlations between the raw scores[3] on the long and short forms were calculated for each scale including the validating scales (Table I).

TABLE I
COMPARISON OF AVAILABLE RETEST RELIABILITY CORRELATIONS WITH CORRELATIONS BETWEEN THE LONG AND SHORT FORMS

Scale	Long vs short correlations (Present study)	Retest reliability correlations (Original authors)
Question	.745	
Lie	.846	
Validity	.927	
Hypochondriasis	.669	.80[a]
Depression	.803	.77[a]
Hysteria	.874	.57[b] and .47[c]
Psychopathic Deviate	.519	.71[b]
Masculinity-Femininity	.756[f]	
Paranoia	.783	
Psychasthenia	.720	.74[d]
Schizophrenia	.891	
Hypomania	.585	.83[e]

[a] Based on the test-retest of 40 normals over intervals of three days to more than a year [11].
[b] Based on the test-retest of 47 normals over intervals of three days to more than a year [16].
[c] Obtained from a test-retest of 98 high school girls over an interval of about one year [16].
[d] Based on test-retest of 47 normals over intervals of three days to more than one year [15].
[e] The number of normals and the interval between test-retest for this reliability coefficient were not supplied [16].
[f] Burton [4] reports the reliability of the masculinity-femininity scale, based on a retest of 34 cases, as .70.

The mean of the weighted scores for each validating and diagnostic scale of both forms were calculated and plotted in a graph (Fig. 1). The degree of dif-

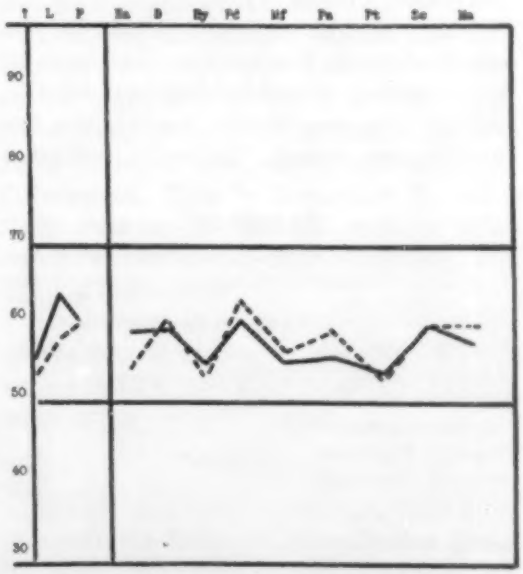

FIG. 1. Mean weighted score profiles of the long (solid line) and short (dotted line) forms.

[3]Correlations were performed with the raw

TABLE II

SIGNIFICANCE OF DIFFERENCES OF MEAN WEIGHTED SCORES FOR THE VARIOUS SCALES ON THE LONG AND SHORT FORMS

Scale	Long Form			Short Form			Differences		
	Mean	SD	σ_M	Mean	SD	σ_M	Mean diff.	$\sigma_{diff.}$	CR
Question	55.2	9.3	.86	53.0	5.5	.50	2.2	.59	3.74
Lie	63.7	10.6	.76	57.3	9.1	.65	6.4	.40	15.88
Validity	58.4	10.5	.52	59.5	11.2	.55	1.1	.21	5.29
Hypochondraiasis	58.0	12.6	1.32	53.8	11.5	1.20	4.2	1.03	4.07
Depression	58.3	14.2	1.15	61.1	15.4	1.25	2.8	.76	3.69
Hysteria	54.8	16.4	1.07	53.4	15.0	.97	1.4	.52	2.69
Psychopathic Deviate	59.8	13.7	1.73	63.2	11.9	1.51	3.4	1.60	2.13
Masculinity-Femininity	55.1	11.7	1.05	56.8	10.8	.97	1.7	.71	2.39
Paranoia	55.7	12.0	.99	58.2	13.1	1.09	2.5	.69	3.62
Psychasthenia	53.8	11.4	1.10	52.9	11.1	1.07	.9	.81	1.11
Schizophhrenia	58.7	13.4	.81	58.8	12.0	.72	.1	.39	.26
Hypomania	56.7	12.4	1.46	58.4	11.1	1.30	1.7	1.26	1.34

ference in mean weighted scores on each scale of both forms and the significance of these differences were also computed (Table II).

The patients were diagnosed by the psychiatric staff after they had been in the hospital approximately one month. Within this one month period, the patients were tested. Table III indicates the various diagnostic categories represented in the group.

TABLE III

DIAGNOSTIC CATEGORIES OF SUBJECTS

Diagnosis	Number
Schizophrenia	11
Manic-Depressive Psychosis	5
Psychopathic Personality	4
Psychoneuroses	3
Organic Psychoses	4
Without Psychosis	3

scores rather than the weighted scores because the distribution of weighted scores did not include all scores that could possibly be attained on the test, e.g., on the depression scale, although the possible range of raw scores is from 0 to 60, the range of raw scores included in the norms are from 8 to 54.

RESULTS

The correlations between the scales of the long and short forms range from .927 to .519. The retest reliability coefficients for the various scales reported by the authors of this inventory in their manual [10] range from .710 to .830 but the coefficient for each specific scale is not indicated. A perusal of the original research on this test, however, reveals retest correlations ranging from as low as .470 to .830 (Table I). It is not clear at present why there exists this discrepancy between the reliability coefficients reported in the manual and those reported earlier in the original research. However, it is of interest that the reliability coefficients obtained in the present study compare favorably with the latter specifically reported reliability data. When the correlations derived from the present study are com-

pared with those reported by the authors in their manual, on only three scales[4] do the correlation coefficients in the former fall below the lowest correlation [.710] quoted by the latter and two[5] fall above the highest.

Examination of the profiles of the mean weighted scores (Fig. 1) shows that the curves very markedly coincide. The greatest difference in mean scores occurs on the lie scale, with the higher score occurring on the long form of this test. Among the diagnostic scales, the greatest difference occurs on the hypochondriasis scale with the highest score on the long form again. However, Table II shows that more than half of the higher mean scores occur on the short form of the test, so that reducing the length of the test has not reduced scores on all scales.

Table II reveals that in one half of the scales statistically significant differences were found between mean weighted scores of the long and short forms. The lie scale showed the greatest degree of significance of difference. On the scales showing statistically significant differences, the higher mean scores of three of the scales[6] were on the long form and three[7] were on the short form.

The average administration and scoring time for the long form was 103 minutes and for the short form 67 minutes (Table IV). This represents a saving of 36 minutes or slightly more than one third the time required to administer and score the long form.

TABLE IV

AVERAGE TIME FOR ADMINISTRATION AND SCORING BOTH FORMS OF MMPI

Average time for	Long form	Short form
Administration	73 minutes	46 minutes
Scoring	30 minutes	21 minutes
TOTAL	103 minutes	67 minutes

DISCUSSION

The range of correlation coefficients between the scales of the long and short forms in this study is greater than that reported by the authors of the inventory in their manual, but compares favorably with the correlation coefficients published for the specific scales. Since the population used in this study is a psychiatric one, as compared to a normal population used by the original authors for their reliability data, considerable variability in test-retest situations could be expected if the long form alone had been used with this population even though caution was exercised in selecting patients who presumably gave representative results. Unlike other areas of measurement, the area of personality testing is subject to greater variability due to the fact that personality itself is highly dynamic and subject to fluctuation.

Although significant statistical differences were reported for half of the scales, inspection of the graph showing mean weighted scores for each scale reveals little, if any, significant clinical differences. This is important to consider because the profile used in this test is the technique that establishes its clinical usefulness.

Since there is no general lowering of scores for all scales on the short form, no correction statistics need be employed in using the short form.

[4]Hypochondriasis, psychopathic deviate, and hypomania scales.

[5]Hysteria and schizophrenia scales.

[6]Question, lie and hypochondriasis scales.

[7]Validity, depression, and paranoia scales.

SUMMARY

Thirty psychiatric patients were given the short form and the original form of the MMPI in order to determine the reliability of the shorter test for clinical practice. More than one-third of the time required for the longer form was

eliminated by using this shortened version. Coefficient correlations of individual scales on test-retest with these two forms range from .519 to .927. Although statistically significant differences were found between the mean weighted scores of half the scales, these results were not clinically significant as judged by profile results.

Received November 18, 1948.

References

1. Abramson, H. A. The effect of alcohol on the personality inventory (Minnesota). *Psychosom. Med.*, 1945, 7, 184-185.
2. Benton, A. L. The Minnesota Multiphasic Personality Inventory in clinical practice. *J. nerv. ment. Dis.*, 1945, 102, 416-420.
3. Benton, A. L. and Probst, K. A. A comparison of psychiatric ratings with Minnesota Multiphasic Inventory scores. *J. abnorm. soc. Psychol.*, 1946, 41, 75-78.
4. Burton, A. The use of the masculinity-femininity scale of the Minnesota Multiphasic Personality Inventory as an aid in the diagnosis of sexual inversion. *J. Psychol.*, 1947, 24, 161-164.
5. Ferguson, R. G. A useful adjunct to Minnesota Multiphasic Personality Inventory scoring and analysis. *J. clin. Psychol.*, 1946, 2, 248-253.
6. Gough, H. D. Diagnostic patterns on the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1946, 2, 23-37.
7. Gough, H. G. Simulated patterns on the Minnesota Multiphasic Personality Inventory. *J. abnorm. soc. Psychol.*, 1947, 42, 215-225.
8. Harmon, L. R. and Wiener, D. W. Use of the Minnesota Multiphasic Personality Inventory in vocational advisement. *J. appl. Psychol.*, 1945, 29, 132-142.
9. Hathaway, S. R. The personality inventory as an aid in the diagnosis of psychopathic inferiors. *J. consult. Psychol.*, 1939, 3, 112-117.
10. Hathaway, S. R. and McKinley, J. C. *Manual for the Minnesota Multiphasic Personality Inventory.* Minneapolis: Univ. of Minnesota Press, 1943.
11. Hathaway, S. R. and McKinley, D. C. A multiphasic personality schedule (Minnesota): III. The measurement of symptomatic depression. *J. Psychol.*, 1942, 14, 73-84.
12. Leverenz, C. W. Minnesota Multiphasic Personality Inventory: An evaluation of its usefulness in the psychiatric service of a station hospital. *War. Med.*, 1943, 4, 618-629.
13. McKinley, J. C. and Hathaway, S. R. The identification and measurement of the psychoneuroses in medical practice: The Minnesota Multiphasic Personality Inventory. *J. Amer. med. Ass.*, 1943, 122, 161-167.
14. McKinley, J. C. and Hathaway, S. R. A multiphasic personality schedule: II. A differential study of hypochondriasis. *J. Psychol.*, 1940, 10, 255-268.
15. McKinley, J. C. and Hathaway, S. R. A multiphasic personality schedule (Minnesota): IV. Psychasthenia. *J. appl. Psychol.*, 1942, 26, 614-624.
16. McKinley, J. C. and Hathaway, S. R. The Minnesota Multiphasic Personality Inventory: V. Hysteria, hypomania and psychopathic deviate. *J. appl. Psychol.*, 1944, 28, 153-174.
17. Meehl, P. E. Profile analysis of the Minnesota Multiphasic Personality Inventory in differential diagnosis. *J. appl. Psychol.*, 1946, 30, 517-524.
18. Meehl, P. E. and Hathaway S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, 30, 525-564.
19. Morris, W. W. A preliminary evaluation of the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1947, 3, 370-374.
20. Schiele, B. C., Baker, A. B., and Hathaway, S. R. The Minnesota Multiphasic Personality Inventory. *Journal-Lancet*, 1943, 63, 292-297.
21. Schmidt, H. O. Test profiles as a diagnostic aid: The Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1945, 29, 115-131.

THE ROSENZWEIG P-F TEST AS A MEASURE OF FRUSTRATION RESPONSE IN SEMISTARVATION

BY JOSEPH C. FRANKLIN[1] AND JOSEF BROZEK
UNIVERSITY OF MINNESOTA

FRUSTRATION, the blocking of the organism in reaching a goal, together with its consequences, is assigned an increasingly critical role in the systematic study of behavior [2]. Ingenious animal experimentation which has contributed much to the clarification of the biodynamics of frustration has stimulated and been paralleled by considerable work at the human level [11, 13, 12, p.6]. In general, the results of these studies confirm the insights gained through clinical experience and support theories, going back to Freud, that early frustration exercises important and profound effects throughout the subsequent life of the individual [6, 21]. The relations between frustration and perception also have been studied [20].

Pioneering in this field, Rosenzweig has developed a "projective" Picture-Frustration test [15] for the measurement of human frustration response.

Rosenzweig [14] differentiates *primary* and *secondary* frustration. According to his classificatory scheme, *primary* frustration is occasioned by deprivation—he uses hunger as the prototype of such frustrations. *Secondary* frustration is instigated by obstructions or events standing in the way of satisfying a need—in this instance, of obtaining food. Protracted semistarvation then should constitute one of the most striking conditions of chronic primary frustration. Moreover, it may be assumed that in such a situation the interaction of primary and secondary frustration responses would tend to reinforce and augment each other.

We had an opportunity to apply this P-F test during a semistarvation-rehabilitation experiment conducted at the Laboratory of Physiological Hygiene, University of Minnesota [7]. The easily administered and relatively objectively scored P-F test was included in a large battery of instruments used in the experiment, the primary purpose of which was to investigate under strictly controlled conditions the relative effectiveness of different types of diets in bringing about recovery from prolonged inanition. The starvation which preceded rehabilitation provided the opportunity to examine the physiological and psychological concomitants and effects of severe undernutrition. Specifically, the conditions of a rigorously controlled dietary stress permitted a study of the validity and sensitivity of Rosenzweig's Picture-Frustration test in a realistic context.

METHOD

The P-F test has been described in detail elsewhere [15]. It consists of 24 roughly-drawn cartoons illustrating common frustrating situations. In each picture two central persons appear. One person is shown saying words which instigate frustration or describe a situation calculated to frustrate the other person. The subject is asked to write down the very first comment or reply which comes to his mind. He is oriented toward giving his immediate response rather than a considered judgment. A

[1]Now at The Johns Hopkins University.

small blank caption-box is provided for the subject's response. The characters are drawn with facial features omitted in order to allow maximum opportunity for "projective" statements. The subject is requested to avoid being humorous and to work as quickly as he can.

In Rosenzweig's terminology, the frustrating situations are of two types. In the *ego-blocking* situations, some obstacle, personal or impersonal, interrupts, disappoints, deprives, or otherwise directly frustrates the subject; e.g., the "frustrating" person comments to the subject, "It's a shame that my car had to break down and make you miss your train." In the *superego-blocking* situations, the "frustrator" makes charges against or incriminates the other person; e.g., in one situation he says to the subject, "You're a liar." According to the hypothesis underlying the scoring and interpretation of the responses the subject may be expected more or less to identify himself with and to use the character for whom he answers as his own mouthpiece.

Subject's responses are scored for the *direction of aggression* (extra-, intro-, and impunitive) and the *type of reaction* (*obstacle-dominant*, *ego-defensive*, and *need-persistent* according to the focus of the response on the *barrier*, as the cause of frustration, on the *ego* of the subject, or on the need for *solution* of the problem). All but the most unusual responses may be scored by combinations of the two criteria, the direction of aggression and the type of reaction. Refinements introduced later into the test provide for analysis of trends in responses and for comparing the pattern of a subject's responses with a group norm (the *group conformity rating*). Rosenzweig's P-F test combines the testing techniques of word association and thematic apperception for the measurement of qualitative and quantitative patterns of response to frustration.

Conditions of the Experiment. Thirty-six "normal" young men between the ages of 20 and 33 (mean age 25.5, SD 3.5 years) served as volunteer subjects. The body build, physical performance capacity, and personality of these men varied within broad limits. Excluding the follow-up studies, the experiment lasted for approximately a year. For the majority of tests control data were obtained during three months of standardization (November 1944 to February 1945) at which time the subjects were maintained on a "good" diet averaging 3,492 calories daily per man. Six months of semistarvation followed (February to July 1945) with the average daily intake reduced to 1,570 calories. The diet was planned to simulate in quality and quantity the food available in western and central Europe under the conditions of severe food shortage during the Second World War. The men lost, on the average, 24 per cent of their prestarvation weight. The semistarvation period was followed by three months of controlled rehabilitation (ending in October 1945). During rehabilitation, four men having been dropped from the experiment, the remaining 32 subjects were divided into four groups of eight men each. These groups were placed on diets averaging 2,377 ("Zero" group), 2,692 ("Low" group), 3,123 ("Good" group), and 3,392 ("Top" group) calories a day, respectively. These values are averages for the whole period of rehabilitation.

The P-F test was administered during the 24th week of semistarvation and, again, after 12 weeks of nutritional rehabilitation at which time the subjects in the four caloric groups, lowest to highest, had regained 21, 30, 41, and 57 per cent, respectively, of weight lost in semistarvation. Although it

would have been more desirable from the point of view of experimental design to have tested the men prior to the semistarvation period rather than use recovery data for control performance, it was not possible for technical reasons to do this. However, it should be noted that at the end of the period of controlled rehabilitation substantial physiological and psychological recovery had taken place in proportion to the level of caloric intake. The recovery was sufficient to admit scores obtained at the end of this period as control data for our purposes.

Scoring. One of the authors (J. C. Franklin) and another staff psychologist (H. Guetzkow), both of whom became thoroughly familiar with the scoring procedure, independently scored the record blanks. Subsequently, each record was jointly re-examined and discrepancies in scoring were discussed. In those cases where disagreements in scoring could be reconciled, suitable revisions in scores were made. Where differences in scoring persisted after joint consideration, the original scores were left unchanged. This procedure was repeated in the handling of the records obtained from the second administration of the test at the end of twelve weeks of rehabilitation.

Combining the two administrations of the test, the following examiner reliability coefficients were obtained for the six scoring categories: extrapunitive .86; intropunitive .87; impunitive .82; obstacle-dominant .66; ego-defensive .55; and need-persistent .93. The overall scoring consistency approximated that reported by Clarke, Rosenzweig and Fleming [1] and was, therefore, regarded as satisfactory. Subjects' scores were taken as the average of scores assigned by each of the two examiners. Reliability coefficients for group conformity ratings, calculated for each administration of the test, were .81 at the end of semistarvation (hereinafter designated as S24) and .89 after 12 weeks of rehabilitation (hereinafter designated as R12).

RESULTS

Semistarvation. The medians and quartiles of semistarvation and rehabilitation data for our experimental subjects are given together with parallel data for Rosenzweig's norm group of medical interns, the more comparable of his two male standardization groups, in Table I. Thus presented the data can be only impressionistically interpreted. At

TABLE I

MEDIANS AND QUARTILES OF PERCENTAGES IN THE VARIOUS SCORING CATEGORIES OF THE P-F TEST FOR STANDARDIZATION AND THE MINNESOTA EXPERIMENTAL GROUPS

	Standardization (N = 50 interns)			Experimental Subjects (N = 32)					
				at S24*			at R12†		
Response Category	Q1	Mdn.	Q3	Q1	Mdn.	Q3	Q1	Mdn.	Q3
Direction of aggression									
Extrapunitive	32	42	48	46	52	60	45	58	69
Intropunitive	21	28	36	23	27	35	19	27	35
Impunitive	24	31	36	17	29	34	19	25	34
Type of reaction									
Obstacle-dominant	13	18	23	24	28	31	23	27	34
Ego-defensive	41	46	56	50	56	59	49	56	62
Need-persistent	27	33	39	16	23	34	17	26	33

*24th week of semistarvation †12th week of rehabilitation

S24 extrapunitive responses of the 32 starved subjects exceeded by approximately 10 per cent the standardization group median percentage. Similar elevation in obstacle-dominant and ego-defensive responses occurred. The frequency of need-persistent responses was markedly lower than for the norm group. Median and quartile values for the remaining two categories were practically identical for both groups.

The statistical differences between the responses of the starved men and of "normal" individuals will be evaluated with reference to the data obtained for the group of 50 medical interns.[2] In Table II the data also show (as in Table I) more frequent extrapunitive, obstacle-dominant, and ego-defensive responses and, further, that the experimental subjects gave significantly fewer impunitive responses. The apparent difference in need-persistent responses between the groups is revealed as insignificant.

Mean group conformity rating at S24 was 71 per cent.

Rehabilitation. Some indication of increased extrapunitiveness is seen in Table I. However, a statistical comparison of S24 and R12 data (Table III) shows no significant differences between the semistarvation and terminal rehabilitation records. Moreover, analysis by caloric subgroups disclosed no relationship between level of refeeding and changes in the distribution of responses in the test categories. Mean group conformity rating dropped to 66 per cent at R12. This decline was not statistically significant.

[2]We are indebted to Dr. Saul Rosenzweig for his cooperation in providing these data. In a personal communication he emphasized the tentativeness of the norms.

TABLE II

DIFFERENCES IN THE NUMBER OF RESPONSES IN THE VARIOUS CATEGORIES OF THE P-F TEST BETWEEN STANDARDIZATION AND THE MINNESOTA EXPERIMENTAL GROUPS AFTER 24 WEEKS OF SEMI-STARVATION

	Number of Responses					
	Standardization		Minn. Subjects		Diff.	
Response Category	(50 interns) M	SD	(N=32) S24 M	SD	S24–Norm M	F
Direction of aggression						
Extrapunitive	9.1	3.2	11.7	3.3	+2.6	12.40**
Intropunitive	6.9	1.9	6.4	2.7	—0.5	0.83
Impunitive	7.1	2.2	5.9	2.2	—1.2	5.80*
Type of reaction						
Obstacle-dominant	4.2	1.7	6.4	1.8	+2.2	30.50**
Ego-defensive	13.5	2.6	11.9	1.5	—1.6	12.45**
Need-persistent	5.4	2.2	5.7	3.0	+0.3	0.23

For 1 and 80 degrees of freedom: $F_{.05} = 3.96$* $F_{.01} = 6.96$**

TABLE III

CHANGES IN THE NUMBER OF RESPONSES IN THE VARIOUS CATEGORIES OF THE P-F TEST BETWEEN S24 AND R12 FOR THE MINNESOTA EXPERIMENTAL GROUP

	S24				R12				Difference S24—R12	
	Mean		S.D.		Mean		S.D.			
Response Category	No.	%	No.	%	No.	%	No.	%	Mean	F
Direction of aggression										
Extrapunitive	11.7	49	3.3	14	12.4	52	3.8	16	+.07	2.12
Intropunitive	6.4	27	2.7	11	6.1	25	2.5	10	—.03	0.89
Impunitive	5.9	24	2.2	9	5.5	23	2.4	10	—0.4	0.84
Type of reaction										
Obstacle-dominant	6.4	27	1.8	7	6.2	26	2.2	9	—.02	0.55
Ego-defensive	11.9	50	1.5	6	12.1	51	2.1	9	+0.2	0.16
Need-persistent	5.7	23	3.0	13	5.7	23	3.0	13	0.0	0.00

For 1 and 31 degrees of freedom: $F_{.05} = 4.16$.

DISCUSSION

In our experiment we imposed a controlled stress situation of chronic semistarvation lasting for six months. During and at the culmination of the stress, clinical observations and the results of physiological and psychological tests registered definite, objective, and marked changes in our subjects, as compared with results obtained by the same methods during the three month period of pre-experimental control. These alterations have been described in detail [5, 7, 8, 9, 10, 19]. Physical weakness, hunger, and the personal limitations occasioned by the experimental regimen were powerful sources of frustration. The process of semistarvation blocked or seriously impaired the subjects' normal patterns of activity. Thresholds to frustration and frustration tolerance declined sharply. The subjects' day-by-day behavior—as compared with control—clearly showed both generalized emotional states (e.g. anger, anxiety) and more specific and directed states or "feelings" (e.g. hostility, jealousy, inferiority), according to Sargent's [18] useful distinction in frustration states.

In rehabilitation the restoration of normal pre-experimental functions generally paralleled the caloric levels of refeeding. By R12 the subjects in the four caloric groups, lowest to highest, had regained 21, 30, 41, and 57 per cent, respectively, of weight lost in semistarvation. At this time hunger and food deprivation were no longer a group problem. The prospects for complete recovery on the basis of substantial recovery already made were bright, and attitudes toward persisting weakness and fatigability, which gradually disappeared as limiting factors, more optimistic.

The results of the Minnesota Multiphasic Personality Inventory are pertinent. At R12 the recovery in *depression* score expressed as a percentage of S24 deterioration amounted to 72.5, 51.3, 44.2, and 9.6 per cent (highest to lowest caloric groups). Similarly, the *hysteria* scores yielded recovery values of 87.5, 66.7, 33.3, and 24.0 per cent. The two higher caloric groups showed a larger recovery than the two lower groups for *hypochondriasis* but the order was not consistently parallel to caloric intake with respective recovery percentages of 64.0, 74.2, 40.5, and 27.9.

Periodically the subjects rated themselves on a series of critical items grouped in the form of a Complaint Inventory; the average recovery was 76, 76, 50, and 55 per cent in the four caloric groups, highest to lowest. Moreover, when the men were asked to estimate all-over percentage of their rehabilitation at R12 the following figures were obtained; 81, 68, 64, and 59. These illustrative findings point out, firstly, the progressive recovery in all subjects, and, secondly, the presence of differential recovery rates with varied levels of caloric refeeding.

Refeeding not only effectively brought about recovery and corresponding return to "normal" behavior but also removed or drastically reduced frustration instigators and responses present during the stress of semistarvation.

The semistarvation frustration response behavior of our subjects as compared with their "normal" frustration response clearly indicated alteration in intensity, direction, and type of frustration response. Accordingly, the P-F test as a measure of frustration response would be expected to distinguish the frustration response status of the group under the contrasting conditions described.

Actually, the results obtained from the P-F test at S24 were duplicated at R12 and no differential changes in scores occurred in relation to caloric

grouping during rehabilitation. Analysis of the statistically insignificant shift of the mean group conformity rating from 71 (S24) to 66 (R12) per cent showed that 21 subjects lost an average of 17 per cent and 11 gained an average of 20 per cent in group conformity rating (Table IV). Individual changes in group conformity ratings were related neither to the level of refeeding nor to any other identifiable variable.

TABLE IV

CHANGES IN P-F TEST EXPERIMENTAL GROUP CONFORMITY RATINGS DURING REHABILITATION

Caloric Group	"Zero"		"Low"		"Good"		"Top"	
Direction of change	−	+	−	+	−	+	−	+
Number of subjects	5	3	5	3	5	3	6	2

In attempting to explain our anomalous results we must refer to and discuss the test itself.

If the test discloses patterns of frustration response which are stable, persistent, and characteristic of a subject, duplication of test-retest results under two frustration-contrasted circumstances would be expected. However, the marked individual changes as indicated by shifts in individual group conformity ratings (see Table IV) do not support this assumption. On the other hand, if the test scores are sensitive indices of a subject's contemporary psychological field with its specific needs and goals, or again, if a subject's responses are a function of both permanent and transient factors, the duplication of S24 results at R12 for the group and indiscriminate (as between levels of rehabilitation at R12) individual changes cannot be accounted for.

Regarding the test's "projective" function, it should be emphasized here that in test administration the instructions given to the subjects followed the standard practice and that the subjects were familiar with tests employing "projective" associative processes.

What are the criteria against which the degree of "projectiveness" of response can be gauged? In the P-F test are responses of a subject to be considered indicative of what his real life behavior would be in those situations presented to him in the form of paper-and-pencil test-item miniatures? Or are his responses to be considered typical of his behavior in life situations in general on the assumption that the test-stimuli are representative of everyday instigators of frustration to which all are exposed? Or, again, are a subject's responses contextually independent of the test-stimuli which elicited them, having unique and telling value for total personality frustration-reponse assessment?

For this test questions of validity, what the test *measures* and *to what extent*, and those of reliability involving as they must both group and individual interitem and intrasubject variability, remain largely unanswered.

Our test results when compared with external criteria and direct evidence of the frustration response status of our subjects, both under control and experimental conditions, are incongruous. On this basis the operation of "projectiveness" of response in the sense employed in the rationale of the test was not demonstrated. Moreover, our data do not support the view, even assuming the test yields objective results, that scores for each subject are sufficiently equivalent "projective" products to permit quantitative interindividual and intergroup comparisons.

Although differences in direction and type of frustration response should distinguish in some measure given individuals or groups under contrasting conditions—in this case, stress and "nonstress" situations — the dimension of

intensity of frustration response may for many purposes be a more useful and sensitive index of frustration response status than either or both direction and type of reaction. Habits and mechanisms of frustration response tend to be characteristic and relatively constant for any given individual throughout a wide range of circumstances. To the unknowns of the P-F test—what is being tested and at what projective level—must be added the *intensity* of the frustration response.

Rosenzweig's [15] two standardization samples differed on occasion from each other by as much as our experimental subjects differed from either of his "normal" standardization groups. For example, in the Rosenzweig group of 50 "normal" men, employees of a large manufacturing plant (ages 22 to 32 years), the median percentage of ego-defense responses was equal to 57 as compared with 46 for his group of 50 medical interns. At S24 our men had a median ego-defense percentage of 56. Therefore, although statistically significant differences in test category responses between the medical interns and our experimental subjects were obtained, both at S24 and R12, it is impossible on the basis of Rosenzweig's *norms* to relate them objectively to any psychologically real deviations from "normal". It is interesting to note that experimental subjects' mean group conformity rating (a measure of "normality") at S24 was higher than that obtained by the "normal" subjects from whose responses the norm for group conformity was derived.

Surprisingly, the strong need-persistent behavior of the subjects was not evidenced in the P-F test scores, which were no higher than those obtained from the norm group. The restrictions, the regimen, the semistarvation diet, in short, all the thwarting conditions of experimental life were imposed and controlled by the experimenters. To them the subjects appealed under all circumstances for any and everything that might mitigate the hardships of the stress and make for relief, recovery, and rehabilitation. The Rosenzweig need-persistent response emphasizes resolution of the frustration tension by removing or altering the frustration through voluntary action or more usually by requesting or demanding the services of a second person, typically the "frustrator." An important component of our subjects' pattern of reaction to the chronic frustrations of the stress was clearly of the need-persistent type. In this respect the situation of the "normal" standardization subjects must have been vastly different from that of our men.

Rosenzweig offers the P-F test as an objective means of "disclosing patterns of responses to everyday stress" [16] and "for . . . investigation of . . . patterns of reaction in situations of stress" [15].

The present study casts doubt on the validity of the Picture-Frustration test as an objective measure of frustration response. Although we are cognizant of the limitations of our material, both in design and in the nature of the experiment, we believe that our findings point up the need for adequate validation and standardization of the P-F test. In addition to establishing the validity and reliability, the effects of "role," "set," and of memory on results in repeated administration of the P-F test need thorough investigation.

SUMMARY

1. Thirty-six men served as volunteer subjects in a semistarvation-rehabilitation experiment conducted at the Laboratory of Physiological Hygiene, University of Minnesota. Twenty-four weeks of semistarvation (S24) pro-

duced striking physiological and psychological alterations of behavior. Severe primary and secondary frustration was induced by hunger, physical weakness, rigors of the regimen, and overall suffering. After twelve weeks of controlled rehabilitation (R12) substantial recovery had been effected.

2. The Rosenzweig Picture-Frustration test, a "projective" method of "objectively" assessing frustration response, was administered at S24 and R12.

3. Significant differences in scores between a norm group of fifty medical interns and the experimental group were obtained at both S24 and R12. These differences, however, cannot be translated into psychologically meaningful differences because of inadequate standardization and questionable validity of the P-F test.

4. For the experimental subjects, group scores for all test categories at S24 were not significantly different from those secured at R12.

5. The mean group conformity rating (a measure of "normality") at both S24 (71 per cent) and at R12 (66 per cent) exceeded that of the group of "normal" medical interns.

6. During controlled rehabilitation subjects were placed in four caloric groups of eight men each, with daily intakes ranging from 2,377 to 3,392 calories. Recovery in weight and in physiological and psychological functions in general paralleled level of caloric intake. P-F test results disclosed no relationship between level of refeeding and changes in either direction or type of response, or in changes in group conformity ratings.

7. The roles of intrinsic personality effects on test performance are indeterminate.

8. The data do not support the assumption that subjects individually or collectively respond comparably "projectively" to the stimulus materials of the test.

9. Measurement of intensity of frustration response may be a more generally useful and sensitive index of frustration response than either or both direction and type of frustration response, and should be provided for in any test of reaction to frustration.

10. Establishment of reliability requires investigation with especial consideration of the effects of "role," "set," and memory, which may be of considerable importance, on scores in repeated administrations of the P-F test.

11. The findings question the validity of the Rosenzweig Picture-Frustration test. Until it more adequately meets the requirements of an objective instrument the Picture-Frustration test should not be used routinely in diagnostic and clinical work but should be restricted to exploratory and experimental uses[3].

Received December 6, 1948.

REFERENCES

1. CLARKE, HELEN J., ROSENZWEIG, S., AND FLEMING, EDITH E. The reliability of the scoring of the Rosenzweig Picture-Frustration Study. *J. clin. Psychol.*, 1947, 3, 364-370.

2. DOLLARD, J., MILLER, N. E., DOOB, L. W., MOWRER, O. H., AND SEARS, R. R. *Frus-*

[3]This paper presents a small portion of the findings of the Minnesota semistarvation-rehabilitation experiment, carried out under the direction of Dr. Ancel Keys. We wish to express our thanks to the senior staff of the Laboratory of Physiological Hygiene, especially Drs. Olaf Mickelson, Austin Henschel, Henry L. Taylor, and Miss Angie M. Sturgeon for their share in this study. Our appreciation is extended to Dr. Harold Guetzkow who was largely responsible for the inclusion of the Rosenzweig Picture-Frustration test in our test battery and for his participation in the administration and scoring of the test. Grateful acknowledgment for his wise counsel and helpful suggestions during preparation of this paper is due Dr. Richard Lazarus of The Johns Hopkins University. Special recognition is due the subjects without whose complete and unreserved cooperation the experiment would have been impossible.

tration and aggression. New Haven: Yale Univ. Press, 1939.

3. ELLIS, A. The validity of personality questionnaires. *Psychol. Bull.*, 1946, 43, 385-440.
4. FOULDS, G. Fustration types among mental defective juvenile delinquents. *Brit. J. Psychol.*, 1945, 36, 29-32.
5. FRANKLIN, J. C., SCHIELE, B. C., BROZEK, J., AND KEYS, A. Observations on human behavior in experimental semistarvation and rehabilitation. *J. clin. Psychol.*, 1948, 4, 28-45.
6. HUNT, J. MCV. The effects of infant feeding-frustration upon adult hoarding in the albino rat. *J. abnorm. soc. Psychol.*, 1941, 36, 338-360.
7. KEYS, A. Human starvation and its consequences. *J. Amer. Diet. Assoc.*, 1946, 22, 582-587.
8. KEYS, A., BROZEK, J., HENSCHEL, A., MICKELSEN, O., AND TAYLOR, H. L. *Experimental starvation in man.* Minneapolis, Minn.: Publ. Lab. Physiol. Hyg., Univ. Minn., 1945.
9. KEYS, A., TAYLOR, H. L., MICKELSEN, O., HENSCHEL, A., AND BROZEK, J. *Rehabilitation following experimental starvation in man.* Minneapolis, Minn.: Publ. Lab. Physiol. Hyg., Univ. Minn., 1946.
10. KEYS, A., MICKELSEN, O., HENSCHEL, A., BROZEK, J., AND TAYLOR, H. L. *Later stages of rehabilitation following experimental starvation in man.* Minneapolis, Minn.: Publ. Lab. Physiol. Hyg., Univ. Minn., 1946.
11. LEVY, D. M. Experiments on the sucking reflex and social behavior of dogs. *Amer. J. Orthopsychiat.*, 1934, 4, 203-224.
12. LEWIN, K. Constructs in psychology and psychological ecology. *Univ. Ia. Stud. Child Welf.*, 1944, 20, 3-29.
13. MASSERMAN, J. H. *Principles of dynamic psychology.* Philadelphia: Saunders, 1946.
14. ROSENZWEIG, S. An outline of frustration theory. In J. McV. Hunt (Ed.) *Personality and the behavior disorders.* New York: Ronald Press, 1944. Vol. 1, Ch. 11, 379-388.
15. ROSENZWEIG, S. The picture association and its application in a study of reactions to frustration. *J. Person.*, 1945, 14, 3-23.
16. ROSENZWEIG, S., CLARKE, HELEN J., GARFIELD, MARJORIE S., AND LEHNDORFF, ANNEMARIE. Scoring samples for the Rosenzweig Picture-Frustration Study. *J. Psychol.*, 1946, 21, 45-72.
17. ROSENZWEIG, S., FLEMING, EDITH E., AND CLARKE, HELEN J. Revised scoring manual for the Rosenzweig Picture-Frustration Study. *J. Psychol.*, 1947, 24, 165-208.
18. SARGENT, S. S. Reaction to frustration—a critique and hypothesis. *Psychol. Rev.*, 1948, 55, 108-114.
19. SCHIELE, B. C. AND BROZEK, J. "Experimental neurosis" resulting from semistarvation in man. *Psychosom. Med.*, 1948, 10, 31-50.
20. SHERIF, M. *An outline of social psychology.* New York: Harper, 1948.
21. WOLF, A. Dynamics of selective inhibition of specific functions in neurosis. *Psychosom. Med.*, 1943, 5, 27-38.

THE GUILFORD-SHNEIDMAN-ZIMMERMAN INTEREST SURVEY

BY J. P. GUILFORD, E. S. SHNEIDMAN, AND W. S. ZIMMERMAN
THE UNIVERSITY OF SOUTHERN CALIFORNIA

THE scope of this article is limited to a brief description of the Guilford - Shneidman - Zimmerman Interest Survey [1], with emphasis upon some of its novel features, an account of its development, and a presentation of some preliminary results.[1]

The GSZ Interest Survey was devised to meet what was felt to be a current need for a new general-purpose interest inventory, based upon the best recent evidence available concerning basic and pertinent interest traits, including factor-analysis information. It is intended for use either singly or as a part of a more comprehensive test battery, by the educational or vocational counselor, the industrial tester, and the clinical psychologist. It is applicable to high-school, college, or adult examinees in general. It may be used either in an individual or in a group setting. It contains 360 briefly worded items and is completed in approximately 45 minutes, there being no time limit. It can be administered either with or without supervision by an examiner.

DESCRIPTION OF THE GSZ INTEREST SURVEY

The Survey can best be described in terms of the categories measured, the dual scores, the special *D* category, the type of item, the selection of items, and the mode of scoring.

Categories Measured. There are 18 special interest traits, paired under nine general interest categories. With this comprehensive scheme it is believed that the various nuances of the different major interests can be included and explored by the test. The 18 special interest traits and the nine general interest categories are listed below:

I. *Artistic*
  1. Appreciative
  2. Expressive

II. *Linguistic*
  3. Appreciative
  4. Expressive

III. *Scientific*
  5. Investigatory
  6. Theoretical

IV. *Mechanical*
  7. Manipulative
  8. Designing

V. *Outdoor*
  9. Natural
  10. Athletic

VI. *Business-Political*
  11. Merchantile
  12. Leadership

VII. *Social Activity*
  13. Persuasive
  14. Gregarious

VIII. *Personal Assistance*
  15. Personal Service
  16. Social Welfare

IX. *Office Work*
  17. Clerical
  18. Numerical

It will be noted that within many of the nine groups, a distinction is made

[1] In this article, portions of the test manual [1], written by the same authors, have been liberally quoted without the use of quotation marks, by permission of the copyright holder.

between appreciative versus expressive aspects; between manipulative or craft types of work versus intellectual and planning activities; or between a more responsible assignment — one requiring leadership—versus a less demanding one. In many instances, a difference between scores in a certain pair of traits will indicate level of aspiration. It is believed important that indications of level of aspiration be permitted to come out within each general sphere of interest. The question of a broadly generalized level of aspiration is still an open one.

Dual Scores. A unique feature of the GSZ Survey is the dual score for each of the 18 special interest traits: a *vocational* and a *hobby* score. This plan takes cognizance of the fact that an individual may like an activity as an avocation or hobby but not as a vocation, or vice versa. It is believed that the availability of these two scores in connection with each of the 18 interest traits provides the test interpreter with valuable additional information about the subject's interests and allows him to judge the subject's genuinely vocational interests with more confidence. It is also important for the counselor to be able to suggest possible satisfying ways for an individual to spend his leisure time as well as to suggest how he may earn a living and like it.

Special "D" Category. In the Survey, there are three possible responses for each item, labeled *D*, *H*, and *V*. *H* and *V* stand for "hobby" and "vocation," respectively, and permit the computation of the dual score for each interest field. In some standard inventories, *D* stands for "dislike." In the GSZ Survey, *D* represents *all* categories not specifically delimited by the *H* and *V* responses; *D* means dislike, disinterested, or don't know. The significance of this is that only the positive responses to *H* and *V* are counted and the ambiguities of the negative responses are eliminated.

Type of Item. Of the different types of items used in interest questionnaires, one kind presents alternative propositions and the examinee is asked to indicate preferences; another type asks the examinee to rate his degree of endorsement of the names of different vocations, school subjects, etc., on a kind of "absolute" scale; a third type of item, found in the GSZ Interest Survey asks for a direct response to various typical activities or aspects of generic interest fields. With the preference type of item, the result is a profile of scores which tend to exaggerate the differences among an individual's interests. The method artificially lowers the correlations between scores, even to the extent of making them spuriously negative. The test interpreter may easily lose sight of the comparative, rather than absolute, meaning of the different scores and may misinterpret the results accordingly.

The use of the second type of item assumes that the examinee's response to a general vocational label indicates that he knows the specific activities within the vocation and to what degree he likes them. General experience would indicate that that is often assuming too much. The examinee's responses, therefore, often do not mean what they seem to mean.

Another practice, that of scoring the same item for more than one trait, also introduces spurious (usually positive) intercorrelation among scores.

The GSZ Interest Survey requires a minimum of discrimination on the part of the examinee; it scores each item for only one trait; and it employs items which are specific enough to sample the actual activities within a generic interest area. These features should reduce the chances of spurious interrelationships, either positive or negative, among

the interest fields, and make the meaning of scores more univocal.

Selection of Items. The GSZ Interest Survey, Form A, was preceded by a preliminary form consisting of 540 items. This was administered to approximately 700 subjects: 400 high-school students, evenly divided as to sex, beginning the 11th grade, and to 300 male college students in a sophomore course. An internal-consistency item analysis was performed on the data from the college group. On the basis of this item analysis, the 20 most internally consistent items in each special interest field were selected to comprise Form A.

Scoring. The scoring method makes possible a rapid computation of the dual (*H* and *V*) scores for the 18 interest fields. An experienced clerk can score papers, each with its 36 scores, at the rate of one every two or three minutes. Items within the test booklet are so rotated that all the responses for any particular interest field lie within a single horizontal row on the answer sheet. Scoring is accomplished by merely counting the number of marked responses in each of the 36 rows, and entering the count at the end of the row.

INTERPRETATION AND DEPENDABILITY OF SCORES

Profile Sheet. For the sake of interpretation of raw scores in terms that are familiar to test users, a profile sheet has been prepared. From distributions of scores obtained from high-school and college students of the two sexes separately, norms have been compiled giving equivalent *C*-scores, centiles, and *T*-scores for all 18 traits, hobby and vocational scores separately. In many instances, the rarity of *H* and *V* responses leads to positively skewed distributions. Discriminations are thus best where it is most important to make them—among individuals with strong interests. The true form of distributions of a population upon these 18 variables is, of course, unknown. We have assumed normality arbitrarily for the sake of providing interpretations in familiar terms.

The high-school norm group was composed of students, of whom 180 were boys and 176 were girls, entering the 11th grade in a Southern California small-city school. The college norm group was composed of students in a sophomore course in a Southern California university, of whom 389 were men and 99 were women. The college men were mostly within the age range from 20 to 25. The college women were mostly within the 18-19 year age groups. These norms are regarded as provisional and are to be revised as results from other parts of the country are obtained.

TABLE I

BRIEF SUMMARY OF ESTIMATES OF RELIABILITY (INTERNAL CONSISTENCY) OF SCORES

	Males		Females	
	Hobby Score	Vocation Score	Hobby Score	Vocation Score
Median	.89	.90	.82	.86
Range	.73-.95	.75-.96	.76-.95	.60-.91

Vocational Implications of Scores. In the manual for the Survey [1], a listing is furnished of the names of the interest traits and under each one some suggestions as to appropriate vocations or positions. Individuals with high standard scores in each trait might be expected to find most of the vocations listed under it to be congenial if not satisfying. The lists are not exhaustive, but they should imply other vocations of similar nature.

Score Reliabilities. The test user properly looks for evidence of reliability of the scores that he uses for predictive purposes. In Table I is given a rough idea of reliabilities of the 36 scores provided by the Survey as based upon ran-

domly selected answer sheets filled out by 100 boys and 100 girls in the high-school sample. First-half scores were correlated with second-half scores and the coefficient corrected by the Spearman-Brown formula. Owing to the skewed distributions, Pearson *r*'s were out of the question for most of the traits. Tetrachoric *r*'s were computed instead, as estimates of Pearson r's, assuming that the population distribution on each interest variable is normal. From these *r*'s no standard error of measurement that would have any practical meaning could be computed. The correlations merely tell us whether one half of the items scored for a trait are psychologically homogeneous with those of the other half. If these were reliability coefficients of the more usual type, based upon computed Pearson *r*'s, one could conclude that they indicate a generally satisfactory degree of internal consistency, considering the fact that each score is based upon only 20 items. It is likely that similar coefficients for college samples would tend to be slightly higher, on the supposition that older and better informed individuals would yield more stable reactions to similar items. The generally wider dispersions of scores (which can be noted on the profile sheets) for college norms can be cited in support of this conclusion.

Intercorrelations. Citations of practical validity data for interest scores are generally lacking for interest inventories, because appropriate external criteria have usually not been available. Production information is generally unsuitable unless samples are extremely large, for correlations with that type of criterion are usually very low. Satisfaction in work, alacrity of effort, and general morale, all difficult to assess, would be more suitable as criteria.

Evidences of internal or factorial validity are to be found in intercorrelation data. On the one hand, one would like to find very low correlations between categories presumed to represent independent interest factors. It is believed that the nine main category scores will show generally low intercorrelations with the possible exceptions of the pair Artistic and Linguistic, and the pair Business-Political and Social Activity. The pairs within main groups are expected to intercorrelate strongly positively but considerable useful discrimination can still probably be achieved.

TABLE II

TETRACHORIC CORRELATIONS BETWEEN HOBBY AND VOCATIONAL SCORES FOR THE 18 SPECIAL INTEREST TRAITS
(N = 300)

Interest Trait	Correlation
Artistic appreciative	.44
Artistic expressive	.53
Linguistic appreciative	.50
Linguistic expressive	.47
Scientific investigatory	.51
Scientific theoretical	.45
Mechanical manipulative	.57
Mechanical designing	.23
Outdoor natural	.43
Outdoor athletic	—.17
Business-political merchantile	.04
Business-political leadership	.20
Social activity persuasive	.42
Social activity gregarious	.04
Personal assistance personal service	.37
Personal assistance social welfare	.44
Office work clerical	.63
Office Work numerical	.59

Correlations between *H* and *V* scores for the same trait might be expected to be generally high. There are results which indicate the contrary. Table II presents some data based upon 300 college men.[2] Tetrachoric *r*'s were computed, assuming normal distributions of

[2]We are greatly indebted to Mr. James Parkinson, Jr. and Mr. Kenneth Foster for the use of the data in Table II.

population interest values and assuming rectilinear regressions. For many of the traits, the correlations tend to be near the value .50; only two are distinctly higher than that figure, and some are even as low as zero, which confirms the usefulness of the dual scores. All such correlations are probably distinctly lower than the reliability coefficients that would correspond to them.

Wherever *H-V* correlations are very low, there is considerable presumption of a genuine hiatus between hobby interest and vocational interest. Two examples are for the traits Outdoor Athletic and Social Activity Gregarious. In either case, the activities proposed are so common in everyday life that many individuals, regardless of their vocational inclinations, are either disposed toward them or not. In general, the results are an emphatic warning that it is often unsafe to depend upon a knowledge of an individual's hobby interests as a basis of assessment of vocational interests. The indications are that liking for hobbies may have some real predictive value in the case of office-work assignments and perhaps also mechanical, scientific, linguistic, and artistic vocations, but too much reliance should not be placed upon such information without better evidence than is now available.

Complete data on intercorrelations of hobby scores and of vocation scores will be available in the near future and will be reported in a supplement to the manual.

SUMMARY

In the Guilford Shneidman-Zimmerman Interest Survey attempts were made to produce a test of comprehensive design that employs recent evidence concerning interest patterns, including factor-analysis information. It is a general-purpose survey, applicable to high-school, college, and general adult subjects. Its characteristics include the measurement of 18 interest traits grouped under nine general-interest categories; the use of dual—hobby and vocation—scores; the special use of the *D* category; the use of a type of item that samples preferences for specific activities directly; and a scoring method that is efficient and rapid. A discussion of the profile sheet, the vocational implications of scores, the score reliabilities and intercorrelations has been presented. Evidence for the hiatus between hobby and vocational interests in many areas has been provided. Additional statistical data will be furnished at an early date.

Received December 30, 1948.

REFERENCE

1. GUILFORD, J. P., SHNEIDMAN, E, S, AND ZIMMERMAN, W. S. *Manual for the Guilford-Shneidman-Zimmerman Interest Survey*, Beverly Hills, Calif.: The Sheridan Supply Company, 1948.

New Books and Tests

BOOKS

ADLER, ALEXANDRA *Guiding human misfits.* (Rev. Ed.) New York: Philosophical Library, 1948. Pp. 114. $2.75.

Adlerian psychology may be somewhat unfashionable, but an examination of this small persuasive book shows its large and often unrecognized influence on many recent dynamic conceptions. The revised edition has an added chapter on post-traumatic neuroses.

DERI, SUSAN *Introduction to the Szondi test.* New York: Grune & Stratton, 1949. Pp. xiv +354. $5.00.

The long-awaited text on the Szondi method contains the instructions for its administration, which are simple, and those for its interpretation, which are not simple at all. Essentially, the test purports to evaluate the strength and control of a person's need-tensions. Discarding Szondi's original theory of genetic determination, Mrs. Deri attempts some theoretical orientation in terms of Lewinian topology. There are chapters on the interpretation of each of the test's four main vectors, and a final section consisting of a small number of case illustrations. The most astonishing aspect of the book is the total absence of evidence attesting to the method's value other than by reference to the experience of the author, and the lack of any reference to research literature beyond Szondi's three books in German. The absence of public evidence for the test's merit is a matter of preference or philosophy, rather than of neglect or ignorance. Mrs. Deri appeals to *private* evidence. The acceptance of her statements is "left to the goodwill of the reader," who should keep his judgment "suspended until he can convince himself about the clinical validity . . . . of the test." But, alas, the personal feeling of validity has been an unhappy guide since Mesmer; most of us prefer more public and communicable testimony. It may be hoped that skillful and hard-headed research workers will not be repelled by the implausibility of the Szondi until it has been given a thorough and public evaluation.

DEUTSCH, ALBERT *The mentally ill in America.* (2nd Ed.) New York: Columbia Univ. Press, 1949. Pp. xxi +555. $5.50.

Deutsch's history of the care and treatment of mental disorders remains an indispensible reference for students of psychiatry and clinical psychology. It is fortunate that so sound and useful a book is highly readable. Students who start it will hardly lay it down until they have acquired a rich sense of the background of the fears and misconceptions that beset public attitudes toward mental illness even today. The greater part of the revision differs little from the original (1937), except for bringing statistics up to date. Terminal chapters on modern treatment and mental hygiene have been revised appropriately. The one entirely new chapter, on psychiatry in World War II, is perhaps too much concerned with apologies for lack of achievement, which is often blamed on the inertia and hostility of military bureaucracy. There are up-to-date decriptions of psychiatric advances in the Veterans Administration and in the Public Health Service.

DEUTSCH, FELIX *Applied psychoanalysis.* New York: Grune & Stratton, 1949. Pp. 244 $3.75.

This little book consists of four essays on limited-objective applications of analytic psychiatry. The first part describes a goal-limited "sector therapy" for brief treatment by means of an interviewing method of "associative anamnesis." It is notable for the verbatim reporting of a complete series of fourteen interviews. The remaining essays are concerned with psychiatry and social welfare, mental health in industry, and family disruption.

FLEMING, C. M. *Adolescence.* New York: International Universities Press, 1949. Pp. vii +262. $4.50.

A facade of light anecdotes that suggest the "popular" book scarcely hides the really wide scholarship and sound interpretation of this small British volume. Its three parts deal with the family, school, and world adjustments of adolescents. Each chapter has a useful bibliography. The book deserves more attention from serious students of adolescence than does many a more pretentious one.

HARING, DOUGLAS G. (Ed.) *Personal character and cultural milieu.* (Rev. Ed.) Syracuse, N. Y.: Syracuse Univ. Press, 1949. Pp. x +670. $5.00.

This volume is "one anthropologist's answer" to the problem of crowded libraries. The thirty-eight papers, selected to supplement other available compilations, are interspersed with comments and additional references. Authors include Boas, Malinowski, Kluckhohn, Bateson, Benedict, Hallowell, Kardiner, Mead, Roheim and others. There is a general bibliography of 33 pages. The erratic typography of the volume, caused by the lithographic reproduction of original articles, impairs its attractiveness but does not harm its sound utility as a student reference.

HAVIGHURST, ROBERT J. AND TABA, HILDA *Adolescent character and personality.* New York: Wiley, 1949. Pp. x + 315. $4.00.

This research study, by the Committee on Human Development of the University of Chicago, again demonstrates that character is an elusive concept. The subjects were 144 adolescents, all those who became 16 years old in 1942 in a small middlewestern city. Character was measured largely by reputation among peers and adults with respect to honesty, loyalty, friendliness, responsibility and moral courage. These reputations were most significantly correlated with school achievement, social class, and intelligence, which is not surprising in view of the values held by the dominant groups of raters. On the whole, the character of the community is better revealed by the study than is the character of any of the individual adolescents.

HOLLINGWORTH, HARRY L. *Psychology and ethics.* New York: Ronald Press, 1949. Pp. iv + 247. $3.50.

Professor Hollingworth has prepared a spritely essay in a field usually avoided by psychologists. In brief, *ought* has many meanings, of which ten are described explicitly. The good life integrates these many *oughts* optimally. The author finds most modern schools of psychology unable to provide a basis for these imperatives. His own systematic psychology, especially in its concept of scope, is more applicable. There is an appendix summarizing an experiment with an ethical insight test.

HURLOCH, ELIZABETH B. *Adolescent development.* New York: McGraw-Hill, 1949. Pp. x +566. $4.50.

A competent textbook on adolescence, heavily documented with research studies and well illustrated by graphs and tables. The physiological changes of puberty, and other body changes, are described more fully than in most texts. There are chapters on emotions, social behavior, friends and leaders, moral concepts, sex, personality, and maturity. A series of chapters on interests, beliefs and attitudes constitute the central emphasis. The omission of anthropological studies of adolescence is exceptional among today's books in this area.

MCNEMAR, QUINN *Psychological statistics* New York: Wiley, 1949. Pp. vii + 364. $4.50.

No statistics textbook is exactly light reading, but McNemar's is exact, concise and clear. Although higher mathematics is not used, he expects the student to have at least the vocabulary of elementary mathematics, and does not hesitate to use phrases such as "one can use a logarithmic transformation of such a variable," without further explanation. Sampling theory is treated with unusual accuracy, and the three chapters on analysis of variance and covariance are probably the best so far written on that topic. The implications for the design of experiments and the treatment of data are especially valuable for the types of experimental problems encountered in clinical psychology.

MELTON, ARTHUR W. (Ed.) *Apparatus tests.* Army Air Forces Aviation Psychology Program, Report No. 4. Washington: U. S. Government Printing Office, 1947 (issued 1949). Pp. xxi + 1056. $3.50.

A distinctive feature of the Air Force's classification testing program is the use of ap-

paratus tests to measure such characteristics as coordination, dexterity, and speed of response. The report gives extensive data on the development, construction, and validation of the eight major tests used for standard classification purposes, and of some fifty experimental tests. The mass testing of thousands of men by complex apparatus tests brought problems never before encountered by psychologists. The calibration procedures used to reduce variance due to apparatus and the test control mechanisms developed to minimize examiners' errors, are therefore of special interest. The report is illustrated by many photographs, working drawings and wiring diagrams. This volume is the last of the nineteen Aviation Psychology Research Reports to appear, completing a series that is probably the most comprehensive accounting of a coordinated research project in applied psychology.

Mons, W. *Principles and practice of the Rorschach test.* Philadelphia: Lippincott, 1949. Pp. 164. $4.00.

A short treatise on the Rorschach method written by a British medical psychologist, based mainly on the approach of Klopfer and Kelley. Some of the interpretative suggestions of Beck, Piotrowski, Harrower, and Hertz are also noted briefly. The author makes a few original contributions to scoring, including *Wv* for inferior "cheap" wholes, and *Mm* for "expressive poses." The final chapter contains three complete protocols, with clinical data on the cases. While it offers little that is novel, the book is a clear and simple text, quite adequate for an introductory course.

Neill, A. S. *The problem family.* New York: Hermitage Press, 1949. Pp. 224. $2.75.

This breezy little book is mostly about Summerhill School, England, of which the author is headmaster, and about sex, in which the author follows Wilhelm Reich. There isn't much about families.

Pease, Katharine. *Machine computation of elementary statistics.* New York: Chartwell House, 1949. Pp. xii + 208. $2.75.

A very clear elementary manual for computations on common models of the Friden, Marchant, and Monroe calculating machines. Part One describes the basic operations of addition, subtraction, multiplication, division, and square root. Part Two gives procedures for calculating the mean, SD, r, percentiles and standard scores. There is an appendix of computing-sheet layouts, including check formulas. The simplicity of description, and the numerous exercises with answers, make the manual suitable for self-instruction.

Reed, Herbert *Education through art.* New York: Pantheon Books, 1949, (Published in Gt. Britain, 1945). Pp. xxiii + 320. $5.50.

That education is self-realization, and that art is therefore education's essential medium, is a thesis as old as Plato. Reed presents the thesis persuasively and interestingly, integrating educational and psychological theories with a mass of first-hand evidence on children's art. Of psychological interest is the author's diagnosis of the child's personality from each painting or drawing, based on Jung's classification of types: introvert-extravert; sensation-feeling-intuition-thinking; and conscious-unconscious. While the conclusions may be regarded skeptically, there is much material for those concerned with the relationships between art productions and personality. The book is illustrated with 116 reproductions of children's drawings and paintings, four in color.

Rosenzweig, Saul, with Kogan, Kate L. *Psychodiagnosis.* New York: Grune & Stratton, 1949. Pp. xii + 380. $5.00.

Not since Well's *Mental tests in clinical practice* (1927) has there been quite so satisfactory an introductory book on psychological diagnosis. Its greatest value is in the thoroughly clinical flavor, with many case studies that illustrate first the use of each test separately, and later, in integration with others. After a brief introduction on clinical psychology as an art and profession, Part Two describes intelligence, deficit, aptitude and interest, questionnaire, and projective methods, with illustrative protocols. Part Three describes four cases at length with full test results, histories, and interpretations. Part Four reviews a number of special problems of the psychodiagnostician —psychotic evasiveness, neurologic damage, somatic complaint, and amnesia, among others. With all its merits, the book is not without faults. Many test descriptions—e. g. 35 lines to describe the Rorschach—are so scanty as to be meaningless without other references. There is question, too, of the value of including very brief and superficial accounts of mechanical and musical aptitude tests. Clinical psycholo-

gists will like this book. Their students will use it widely, both as a framework for beginning study in psychodiagnosis, and as an integrative summary for more advanced instruction.

ROTMAN, DAVID B. *Addresses and papers, 1934-1948.* Selected and edited with a foreword by Agnes A. Sharp. Chicago: Psychiatric Institute of the Municipal Court, 1948. Pp. xiv + 205.

A memorial volume to Dr. David B. Rotman, who died in 1948 at the age of 54, after twenty years of service as psychiatrist and director of the Psychiatric Institute of the Municipal Court of Chicago. Twenty-eight of Dr. Rotman's papers are included, dealing with psychiatry in the courts, crime and mental disease, alcoholism, and the clinic approach. Some of these articles had appeared in journals, others were addresses to groups of laymen. Throughout, the volume reflects the warm and vital personality of Dr. Rotman, and the esteem that he evoked from his co-workers.

SNYGG, DONALD AND COMBS, ARTHUR W. *Individual behavior, a new frame of reference for psychology.* New York: Harper, 1949. Pp. ix + 386. $3.50.

The pendulum of psychological theory has swung far from the objectivism so neatly defined by the title of Max Meyer's *The psychology of the other one* in 1921. Much of today's most stimulating thinking, especially in the contexts of social and clinical psychology, has become the psychology of *the experiencing one*, variously styled personalistic, self, or phenomenological psychology. In this area Snygg and Combs have made the neatest formulation so far achieved. Their central thesis is that all behavior is determined by the present phenomenal field of the behaving organism—not by the "situation," but by the way he perceives the situation. Traditional psychological problems are handled in that frame of reference. Action, motivation, learning, threat, conflict, readjustment, and even the unconscious are susceptible to restatement in phenomenological terms. Applications to social psychology, to education and to psychotherapy are discussed at length. The book is one that no general or clinical psychologist interested in serious approaches to theory can ignore.

STOUFFER, S. A., SUCHMAN, E. A., DEVINNEY, L. C., STAR, S. A., AND WILLIAMS, R. M., JR. *The American soldier. Vol. I, Adjustment during army life.* Pp. xii + 599.

STOUFFER, S. A., LUMSDAINE, A. A., LUMSDAINE, M. H., WILLIAMS, R. M., JR., SMITH, M. B., JANIS, I. L., STAR, S. A., AND COTTRELL, L. S., JR. *The American soldier. Vol. II, Combat and its aftermath.* Pp. 675. Princeton, N. J.: Princeton Univ. Press, 1949. Two volumes $13.50, each $7.50.

Social psychology's most extensive program of applied research was that of the Research Branch, Information and Education Division of the War Department, in World War II. The comprehensive report of its methods and findings fills these two clear, literate and handsome volumes, published under the editorial auspices of the Social Science Research Council. The reports contain little discussion of methods as such—interviewing, questionnaire, or sampling techniques—but are concerned mainly with findings. Of obvious and direct military significance are studies such as those of attitudes toward war aims and toward combat, of replacement and rotated troops, and of the outstanding practical triumph of the Research Branch, the "point system" for discharge weighted according to the men's own attitudes. More widely applicable results include studies of adjustment to rapid social changes, studies of leadership, and studies of the attitudes toward and of the Negro soldier. Clinical psychologists will find special interest in the surveys of the control of fear, and of psychoneurotic symptoms in the army.

SYMONDS, PERCIVAL M. *Adolescent fantasy.* New York: Columbia Univ. Press, 1949. Pp. xii + 397, 42 plates. $6.00.

Symonds' work is a skillful blend of objective and dynamic approaches, applied to the development of a picture-story technique for exploring the personalities of adolescents. After a review of the literature of the thematic-apperception type of test, there is an account of the development and selection of the pictures used, analyses and inventories of themes, norms, associations to stories, and quantitative relationships. Several chapters are of general value in giving insight into adolescent needs, stresses and fantasies, apart from the particulars of the test itself. There are two long case studies, of a maladjusted boy and of a well adjusted one. The volume is an indispensible background and aid in the use of the Symonds

Picture-Story Test (see *J. consult. Psychol.*, 1949, 13, 67-78).

SYMONDS, PERCIVAL M. *Dynamic psychology.* New York: Appleton-Century-Crofts, 1949. Pp. vii + 413. $3.75.

Dynamic psychology is an undergraduate textbook in psychoanalytic theory, a condensation of the author's earlier *Dynamics of human adjustment.* In the main, the abbreviation has been achieved by the omission of sections and paragraphs, and by the rewriting of some of the more difficult passages. Six chapters of the longer book disappear altogether; in their place are a new introduction and a new chapter on the ego and the self. There are questions and a brief list of recommended readings at the end of each chapter. With its clear style and numerous applications to everyday situations, it is probably the best text to present the psychoanalytic orientation to the beginner.

TUTTLE, HAROLD S. *Dynamic psychology and conduct.* New York: Harper, 1949. Pp. xiv + 448. $3.50.

Mainly a practical book on school management and discipline, this volume suffers from shortcomings all too common to the field of so-called character education. It carries a heavy burden of conventional morality (e.g. "preventing one from becoming enslaved to habits of smoking and drinking"), and a light load of psychological theory (there is much confusion between "conditioning" and learning by "effect").

WILLIAMSON, E. G. (Ed.) *Trends in student personnel work.* Minneapolis, Minn.: Univ. of Minnesota Press, 1949. Pp. x + 417. $5.00.

The development of student personnel work over a quarter of a century is reviewed by forty-three papers presented in November, 1947, at a conference honoring the twenty-fifth year of service by Professor Donald G. Paterson. While so large a number of intentionally diverse papers cannot be integrated briefly, some trends can be noted toward greater consciousness of broad social issues, and toward emphasis on individual counseling and mental hygiene. The volume is an essential resource for those with college personnel responsibilities.

TESTS

HILDRETH, GERTRUDE H. AND GRIFFITHS, NELLIE L. *Metropolitan Readiness Tests.* Grades Kgn.-1. 2 forms, R & S. (60) min. Test booklets ($2.10 per 25) with manual (pp. 32), key, norms, class record. Yonkers, N. Y.: World Book Co., 1949.

A new test of readiness for first grade learning, yielding a reading readiness score, a number readiness score, and a total score. The first four subtests measure comprehension of words and sentences, and visual perception, responses being made by marking pictures. The fifth subtest measures number concepts, and the sixth is a copying test involving visual perception and motor coordination. The draw-a-man test is included as an optional supplement, and may be scored approximately by a five-step comparison scale, or in the more detailed way described by Goodenough. Reliabilities obtained from alternate-form administration are .83 for the reading and number parts and .89 for the total. Validities indicated by correlation with a primary achievement test after five months in the first grade are .53 for the prediction of reading achievement, and .63 for number. Norms are based on 15,000 children in 26 states. The test can be administered to groups of 15, and is attractive and convenient for both children and examiner.

THOMPSON, CHARLES E. *Thematic Apperception Test—Thompson modification.* Adolescent-adult. 1 new form, for Negro subjects. 30 plates ($5.00). Manual, pp. 12 (50¢). Cambridge, Mass.: Harvard Univ. Press, 1949.

Clinical experience in administering TAT to Negro examinees has often shown a paucity of production, due to the difficulty experienced by the subjects in identifying with the white figures in the pictures. In Thompson's revision, 22 of the original pictures have been redrawn to introduce Negro figures in place of white; the blank card is gray instead of white; 7 sufficiently ambiguous pictures required no change; one picture (No. 10) is omitted. The author's clinical impression is that richer records are secured from Negro subjects by the use of the new cards. A further suggestion might be the use of the two sets of cards to explore racial attitudes.

AVAILABLE BACK ISSUES

JOURNAL OF CONSULTING PSYCHOLOGY

The issues listed below are available for sale

YEAR	VOL.	NUMBERS						PRICE PER NUMBER	PRICE PER VOLUME
		JAN. FEB.	MAR. APR.	MAY JUNE	JULY AUG.	SEPT. OCT.	NOV. DEC.		
1937	1	1	2	3	4	5	6	$.60	$3.00
1938	2	1	2	3	4	5	6	$.60	$3.00
1939	3	1	2	3	4	5	6	$.60	$3.00
1940	4	1	2	3	4	5	6	$.60	$3.00
1941	5	1	2	3	4	5	6	$.60	$3.00
1942	6	1	2	3	–	–	6	$.60	$2.40
1943	7	1	2	3	4	5	6	$.60	$3.00
1944	8	1	2	3	4	5	6	$.60	$3.00
1945	9	–	2	3	4	5	6	$.60	$3.00
1946	10	1	2	3	4	5	6	$.60	$3.00
1947	11	–	–	3	4	5	6	$.60	$2.40
1948	12	1	2	3	4	5	6	$1.00	$5.00
1949	13	By Subscription, $5.00						$1.00	

Price for Volumes 1 through 12 $36.80

Table based on inventory of January 3, 1949

Information about the Journal of Consulting Psychology: all numbers but 4 are still available. Number 4 of Volume 6 will not be sold separately, but only with orders for several volumes.

Information about prices: the Journal of Consulting Psychology had the uniform price of $3.00 per volume and $.60 per issue from 1937 to 1947. It is now priced at $5.00 per volume and $1.00 per issue. For foreign orders, $.25 per volume should be added. Address orders to:

AMERICAN PSYCHOLOGICAL ASSOCIATION
1515 Massachusetts Ave. N. W.
Washington 5, D. C.

Outstanding McGRAW-HILL *Books*

FRUSTRATION

The Study of Behavior without a Goal

By NORMAN R. F. MAIER, University of Michigan. *McGraw-Hill Publications in Psychology.* 264 pages, $3.50.

This important work is a report of an experimental program of research on frustration conducted in the author's laboratory during the past ten years. The volume presents the evidence which leads to a new interpretation of behavior expressed when an individual is in the state of frustration. The author departs from the usual approach to abnormal behavior, which seeks to discover how the behavior solves a problem for the individual. Instead, the approach in this book is to indicate that searching for the solution of a problem in the symptom is misleading; rather, according to Professor Maier, symptoms are determined by the principle of availability and are not determined by the end they serve.

MENTAL HYGIENE IN PUBLIC HEALTH

By PAUL V. LEMKAU, School of Hygiene and Public Health, The Johns Hopkins University. *McGraw-Hill Series in Health Science.* 396 pages, $4.50.

This is the first book of its kind in preventive psychiatry to treat the problem of mental health in terms of public health action. Dr. Lemkau traces the development of the personality, illustrating at each stage and change in development the possibilities for modification in the direction of better mental health. An effective basis for setting up constructive programs is presented.

PERSONALITY MALADJUSTMENTS AND MENTAL HYGIENE

A Text book for Students of Mental Hygiene, Psychology, Education, Sociology, and Counseling. New second edition

By J. E. WALLACE WALLIN, Upsala College. *McGraw-Hill Publications in Psychology.* In press.

Here is a revision of an established text which presents a discussion of the outstanding types of personality maladjustments, together with suggestions for overcoming mental conflicts and developing normal personalities. Whole chapters are now devoted to a discussion of the educational elements of the mental hygiene program; psychotherapeutic methods; and the practical mental hygiene treatment of inferiority feelings, compensatory maladjustments, stage fright, and fears.

Send for copies on approval

McGRAW-HILL BOOK COMPANY, Inc.

330 West 42nd Street **New York 18, N. Y.**